Women Also Know Stuff

Challenging the Gender Gap in Political Sophistication*

Abstract

This article 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 $N \approx 35,000$). 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

Political sophistication is a foundational concept in the study of political attitudes and behavior—a crucial moderator impacting a range of mechanisms such as political decision-making and vote choice (Macdonald, Rabinowitz, and Listhaug, 1995; Lau and Redlawsk, 2001), persuasion and motivated reasoning (Zaller, 1992; Lodge and Taber, 2013), or the susceptibility to misinformation (Vegetti and Mancosu, 2020). Yet, fundamental concerns regarding the measurement of political sophistication continue to plague the discipline (Mondak, 2001; Sturgis, Allum, and Smith, 2008; Bullock and Rader, 2021). Scholars usually rely on survey questions that assess people's ability to recall basic facts about political institutions and officeholders as a proxy for sophistication (Delli Carpini and Keeter, 1993; Barabas et al., 2014). In principle, these factual knowledge questions should cover information that is necessary for citizens to make competent decisions in a given context, but determining such a set of items proves to be extremely difficult (Lupia, 2006). 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, different sets of knowledge questions vary in difficulty across subgroups of the population, which can introduce systematic measurement error (Pietryka and MacIntosh, 2013).

One manifestation of such systematic measurement error is the oft-cited gender gap in political 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; Fraile, 2014a). To a certain extent, these findings may reflect genuine differences in political interest between men and women due to gendered socialization (Bos et al., 2021). At least part of the observed gender gap, however, can be attributed to measurement. For instance, men are more willing to guess when answering recall questions, which can inflate their estimated knowledge levels (Mondak and Anderson, 2004; Fortin-Rittberger, 2020). Other research finds that gender differences are attenuated when focusing on gender-relevant political knowledge (Dolan, 2011), by providing policy-specific information (Jerit and Barabas, 2017), or in contexts with more

equitable representation of women (Pereira, 2019; Kraft and Dolan, 2022b).

In this paper, I re-examine the gender gap by proposing *discursive sophistication*—a new measure that is based on how people discuss their political preferences in open-ended responses. Specifically, I develop a framework to assess whether beliefs and attitudes on a range of political issues 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). Furthermore, applied researchers can directly implement the proposed method using a software package available for the statistical programming environment R.[1]

I validate the measure across multiple representative surveys in the United States and Europe encompassing four languages (total $N \approx 35,000$) 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 and factual knowledge are independent predictors of turnout, political engagement, and various manifestations of political competence—suggesting that both measures can be viewed as complements that capture different aspects of political sophistication. 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. Furthermore, I present suggestive evidence that this divergence can be

---

[1]R package available here: ███████████████████████████████.

2

explained by the fact that open-ended responses allow women to focus on different issues than men. In sum, the results suggest that exploring open-ended responses provides new opportunities to examine political sophistication across time and contexts.

Political Sophistication and Factual Knowledge

Public opinion researchers routinely incorporate political sophistication in their empirical analysis—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.[2] For instance, Delli Carpini and Keeter (1993)—a canonical article proposing such a battery—has been cited more than 1000 times since its publication. In short, political knowledge remains a concept of intense scholarly interest and it is frequently measured using standard off-the-shelf recall questions.

The ubiquity of basic recall questions in public opinion research is accompanied by the frequent findings that many people know too little about politics (Delli Carpini and Keeter, 1996; Barabas et al., 2014) and that the discrepancies in information levels can result in unequal representation in the political system (Althaus, 1998; Kuklinski et al., 2000; Gilens, 2001). The underlying reason why scholars focus on people's ability to recall factual information about politics is that these items "more directly than any of the alternative measures, capture what has actually gotten into peoples minds" (Zaller 1992, 21; see also Zaller 1991; Gomez and Wilson 2001). However, there is some reason to doubt this assertion, both from theoretical as well as methodological perspectives.

First, the discipline's exclusive focus on factual political knowledge has been criticized on theoretical grounds. Most importantly, recalling facts about political institutions has little relevance for citizen competence (Lupia, 2006; Cramer and Toff, 2017). Given

---

[2]To name but one example, the American National Election Study routinely asks questions such as, "Do you happen to know which party currently has the most members in the U.S. Senate?"

that there is usually no consensus about what information is necessary in the first place, Druckman (2014) proposes abandoning recall questions as measures of "quality opinion." Instead, the author advocates "*less* focus on the *content/substance* of opinions [...] and *more* on the *process* and specifically the *motivation* that underlies the formation of those opinions" (2014, 478, emphasis in the original). The key distinction should therefore be how citizens approach a political issue and whether they are motivated to engage in elaborate reasoning to arrive at their particular decision.

It turns out that such competent decision-making does not necessarily require citizens to hold large swaths of political information in their declarative memory (i.e., what is being measured by conventional knowledge scales). In fact, people can often use heuristics to navigate the realm of politics without having to remember encyclopedic facts about institutions or actors (Lupia, 1994). Even simple visual cues have been shown to increase political knowledge levels (Prior, 2014). While other research suggests that heuristics do require a baseline level of expertise (Lau and Redlawsk, 2001; Lau, Andersen, and Redlawsk, 2008) and that their effectiveness can depend on the political context (Dancey and Sheagley, 2013), this body of literature shows that competence cannot be reduced to the capacity of citizens to remember facts alone. It is more important that people possess the skills and resources to find the information required in a specific context (e.g., Bernhard and Freeder, 2020). In other words, procedural memory appears to be more integral to political competence than declarative memory (Prior and Lupia, 2008). In a similar vein, Luskin (1990) suggests that individual motivation and abilities help explain political sophistication more so than the availability of factual information.

Beyond these theoretical critiques, there are several methodological issues that cast doubt on the validity of factual knowledge scores as a measure of political sophistication. One problem frequently discussed in the literature revolves around the question whether or not to offer "don't know" options in multiple choice recall questions (Mondak, 2000; Mondak and Davis, 2001; Miller and Orr, 2008). Including such an option can lead to bi-

ased estimates of information levels because they are confounded by people's differential propensity to guess instead of admitting not to know the correct answer (but see Luskin and Bullock, 2011). Other scholars criticized open-ended factual knowledge questions due to problematic coding rules, which do not capture partial knowledge (Mondak, 2001; Krosnick et al., 2008; Gibson and Caldeira, 2009; DeBell, 2013). However, closed-ended recall questions are not without issues either. Conventional item batteries differ with regard to the temporal and topical dimensions of the underlying information—which can have important implications for researcher's conclusions about the antecedents and consequences of political knowledge (Barabas et al., 2014). In addition to question content, recent research reveals how their format (e.g., true-false vs. multiple choice) can further exacerbate assessed knowledge inequalities in society (Fraile and Fortin-Rittberger, 2020).

The increasing reliance on online surveys creates additional concerns for recall questions due to people's tendency to look up answers (Clifford and Jerit, 2016; Höhne et al., 2020). This is particularly problematic if there are systematic differences between respondents' likelihood to cheat when answering knowledge questions (Style and Jerit, 2020). Even if cheating was not an issue in online surveys—because respondents are effectively discouraged from searching for correct answers—factual knowledge scores can still suffer from differential item functioning, since individual recall questions have varying measurement properties across the population (Pietryka and MacIntosh, 2013). Item batteries that are easier to answer for certain groups can therefore exacerbate observed differences in political knowledge—for example between racial groups (Abrajano, 2014).

Please Mind the Gender Gap

Survey researchers not only find that people are not sufficiently informed as a whole, they also attest that women are systematically less knowledgeable than men. For instance, women routinely score lower on political information, interest, and efficacy, which de-

creases 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, Verba, Burns, and Schlozman (1997, 1070) 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." While differences between women and men in political *interest* can certainly be attributed to gendered political socialization (Bos et al., 2021; Wolak, 2020), at least part of the disparities in *knowledge* 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 these systematic differences in the propensity to guess mitigates the gender gap in knowledge but does not eliminate it completely (see also Lizotte and Sidman, 2009; Ferrín, Fraile, and García-Albacete, 2017). Furthermore, recent research further suggests that open-ended question formats that discourage guessing may diminish observed gender differences altogether (Ferrín, Fraile, and García-Albacete, 2018).

Other aspects of the survey context have been shown to affect gender differences in political knowledge as well. McGlone, Aronson, and Kobrynowicz (2006) 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*. Focusing on gender-relevant political knowledge items such as information about women's representation in the federal government has been shown to close—or at least reduce—the gap

(Graber, 2001; Dolan, 2011; Fraile, 2014b; Barabas et al., 2014). Similarly, the gender gap shrinks or disappears when people are asked about specific policies and/or long-standing facts rather than current events (Ferrín, Fraile, and García-Albacete, 2018), practical issues related to the government such as the availability of benefits and services (Stolle and Gidengil, 2010), or in political contexts characterized by more equitable representation of women (Pereira, 2019; McAllister, 2019; Wolak and Juenke, 2021). Importantly, women's lower factual knowledge scores can be easily ameliorated by providing additional information (Jerit and Barabas, 2017) and they do not appear to impede on their political competence. In fact, Dassonneville et al. (2020) find that women are no less likely to vote for candidates who represent their preferences, and are therefore able to participate in politics just as effectively as men.

Overall, the gender gap appears 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.

Back to the Roots: The Structure of Belief Systems

Despite the discipline's reliance on off-the-shelf item batteries, factual knowledge about political institutions has little relevance for competent decision-making in politics, which lead some scholars to suggest that we should start considering alternatives to these types of recall questions (Druckman, 2014). From a theoretical perspective, knowledge scores are all but a proxy for an underlying latent trait—political sophistication—which is usually conceptualized based on people's belief systems instead of focusing on isolated pieces

of factual information stored in declarative memory. 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).

Political sophistication can then be characterized by how these ideas and attitudes (or considerations) are structured along three different dimensions (Luskin, 1987). The first, and most obvious one, is the *size* of a belief system, which simply describes the number of distinct considerations that are available for retrieval. Politics, however, is comprised of a diverse set of independent domains—with some people having a deep grasp of a narrow field and others having a broad and potentially more shallow understanding of various issues. Thus, the second dimension describes the *range* of a belief system across domains (e.g,. different policy issues or other evaluative categories). The last dimension is a belief system's *constraint*, which describes the extent to which considerations are organized in a meaningful way through differentiation and integration of competing cognitions (Luskin, 1987). In other words, this dimension captures whether available considerations are perceived as operating in isolation or are rather as part of a more complex interconnected system, for example by identifying inherent value conflicts (Tetlock, 1983, 1993). To summarize, I conceptualize political sophistication based on the *structure* of individual belief systems along the following three dimensions:

1. **Size:** *The number of considerations associated with a given category or issue.*

2. **Range:** *The dispersion of considerations across different categories or issues.*

3. **Constraint:** *The extent to which considerations are interconnected in a meaningful way.*

Political sophistication, in turn, is 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). Similarly, Tetlock (1983, 1993) coined the term *integrative complexity* to describe the degree to which considerations

related to an issue are interconnected. In short, sophisticated political reasoning should reflect this notion of complex belief systems.

To what extent does political sophistication defined as a complex system of beliefs ultimately facilitate citizen competence? As discussed above, conventional knowledge questions have been criticized because the required information has little relevance for people's ability to make high-quality decisions (Lupia, 2006). As Cramer and Toff eloquently summarize, conventional measures implicitly focus on "what people do *not* know" (2017, 756, emphasis added) by presupposing pieces of information as necessary for political competence. Examining people's political beliefs system, on the other hand, allows us to shift the focus back to *what they do know* and how they use that information. After all, a large, wide-ranging, and highly constrained system of beliefs will help citizens to locate their own interests within the political system, understand the functioning of institutions, assess the performance of the incumbent government, and evaluate the actions of the main political actors (e.g., Converse, 1964).[3]

Measuring Discursive Sophistication

Given that recall questions are only an imperfect measure for political sophistication, it is worth considering alternative—and potentially more imminent—observable implications of the underlying latent trait of interest: complex and highly constrained political belief systems. In the following, I propose a framework that leverages the content of open-ended responses in conjunction with the survey structure to evaluate how people discuss their political beliefs and preferences in their own words (see also Kraft, 2018, for a related analysis of moral reasoning in open-ended responses). To illustrate my approach in the context of a concrete example, consider a questionnaire that asks respondents to answer

---

[3]It should be no surprise that Converse (1964) and others examined open-ended responses in their early 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.

the following open-ended item:

> On the issue of **gun legislation**, please outline the main arguments that come
> to mind *in favor and against* background checks for all gun sales, including at
> gun shows and over the Internet.

Now suppose that this questionnaire includes a whole set of similar prompts on other topics such as **abortion**, **immigration**, **health cure**, and **trade policies**—each asking respondents for both positive or negative considerations related to specific policy proposals. How would a complex and constrained set of political beliefs manifest itself across such a battery of open-ended responses? I argue that each dimension outlined above has direct observable implications for individual response behavior.

First, the *size* of a belief system is defined as the number of available considerations associated with a given category or issue. In the context of open-ended survey questions, a large belief system should therefore allow people to discuss their views by raising a larger number of distinct topics in response to each query. While this could also be achieved through manual coding, 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).[4] Let $\mathcal{W}_i$ denote the set of words contained in a response of individual i. Each word $w \in \mathcal{W}_i$ is assigned to a topic $t^* \in \{1, ..., T\}$, such that $P(t^*|w, X_i) > P(t|w, X_i) \forall t \neq t^*$.[5] 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

---

[4]Please refer to the appendix for additional information. Specifically, see Appendix A for a data overview and Appendix B for descriptive information on open-ended responses, pre-processing, and modeling choices for the structural topic models. Appendix C contains additional robustness checks including a preText analysis to explore sensitivity for alternative model specifications (Denny and Spirling, 2018).

[5]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).

considerations can be written as:

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

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, the *range* of a belief system is defined as the dispersion of considerations across categories or issues. Given a set of survey prompts covering various political issues, high levels of sophistication should correspond with people's ability to respond to each query with comparable levels of elaboration. I therefore quantify the consistency in response behavior across items by computing the Shannon entropy in open-ended response lengths:

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

where p_{ij} is the proportion of words in the response of individual i to question $j \in \{1, ..., J\}$ 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 addresses the level of *constraint* between considerations. The extent to which considerations are interconnected in a meaningful way should be associated with people's ability to differentiate and/or integrate them in their reasoning (Tetlock, 1993). Following Tausczik and Pennebaker (2010), I rely on specific function words as linguistic markers for these processes. More specifically, differentiating competing considerations in speech is usually accomplished using exclusive words (e.g., but, without), while integrating multiple thoughts is accomplished by the use of conjunctions (e.g., and, also). Thus, I measure relative constraint by identifying the number of conjunctions (CONJ$_i$) and exclusive words (EXCL$_i$) in each open-ended response using the Linguistic Inquiry

and Word Count (LIWC) dictionary (Pennebaker et al., 2015):

$$\text{constraint}_i = \frac{\text{CONJ}_i + \text{EXCL}_i}{\max\left[\text{CONJ}_i + \text{EXCL}_i\right]} \tag{3}$$

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.

Together, the three measures can be combined in an additive scale of *discursive sophistication* in political attitude expression:

$$\text{discursive sophistication}_i = \text{size}_i + \text{range}_i + \text{constraint}_i. \tag{4}$$

Overall, a highly sophisticated individual should therefore give a more *elaborate* response across the full *range* of questions by *integrating* and/or *differentiating* multiple considerations. Given each individual input, the resulting metric has a theoretical range between 0 and 3. In order to allow for easier comparisons with conventional additive knowledge scores, I rescale discursive sophistication to mean zero and unit variance. Note that this simple framework makes no assumptions about the direction of people's attitudes or their specific ideology. Crucially, since it is solely based on *how* individuals discuss their preferences, it can be directly applied in various settings to target specific political issues or tasks such as choosing between candidates running for election. In other words, we can study discursive sophistication in well-defined (and potentially narrow) areas by using open-ended questions that focus on attitudes and beliefs that are relevant for a specific context. Researchers interested in citizen competence in local politics, for instance, could field a battery of open-ended questions examining relevant topics such as schooling, zoning, or other areas of local administration.

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; elite 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. 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.[6] After all, being hard to comprehend does not necessarily imply that someone put a lot of thought into a statement.

Data and Analytical Strategy

To evaluate my proposed measure of discursive sophistication, I included the battery of open-ended questions described above in a 2018 wave of the Cooperative Election Study (CES),[7] which consists of a national stratified sample of 1,000 respondents. In addition, I illustrate the versatility and robustness of the approach by applying the measure across multiple previously collected surveys that employ a range of alternative open-ended

---

[6]In fact, pure linguistic complexity is arguably driven more by other factors such as a person's general verbosity or linguistic prowess and therefore less valid as a measure of political sophistication.

[7]Formerly Cooperative Congressional Election Study (CCES).

items. Below is a summary of all data sets and items used in the subsequent analysis:[8]

- **Cooperative Election Study (CES 2018):** 10 open-ended questions targeting policy preferences on gun legislation, abortion, immigration, health care, and trade.

- **American National Election Study (ANES 2020, 2016, 2012):** 8 open-ended likes-dislikes questions targeting preferences for parties and candidates.

- **YouGov Survey (2015)**: 4 open-ended questions targeting policy preferences on gun legislation and health care.

- **Swiss Referendum Surveys (2012-2008)**: 2 open-ended questions asking respondents to justify their vote choice in various policy referenda. These surveys were conducted in three languages (French, German, and Italian).

I proceed by providing descriptive evidence regarding the face validity of discursive sophistication. Next, I assess its construct validity by comparing it to factual knowledge as a predictor of various relevant outcomes such as political participation and engagement. The last validation step consists of comparing discursive sophistication to manually coded levels of justification in open-ended responses. Each of these steps leverage different subsets of the studies listed above, depending on the availability of necessary items. After validating the measure, I assess gender gaps in discursive sophistication and factual knowledge using the complete set of surveys.

A First Look at Discursive Sophistication

While each dimension of discursive sophistication outlined above provides a unique source of variance to the underlying concept (Luskin, 1987), all three are positively correlated.[9] Furthermore, exploratory factor analyses confirm that they load on a single factor with all loadings exceeding 0.5 across the CES and ANES data—thus confirming that we can rely on an additive score to measure discursive sophistication (see Table 1).[10]

---

[8]A detailed description of each data set and the specific question wording is included in Appendix A.

[9]See Appendix B.III for correlation matrices between individual components.

[10]I rely on the CES and ANES here since these surveys employ a larger set of open-ended questions.

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

How does this discursive sophistication score compare to alternative metrics of political knowledge? As discussed, the standard approach to measuring political knowledge in surveys is to ask a set of factual questions about political institutions. The CES and ANES include such a set of basic recall items, inquiring for example about presidential term limits or the majority party in either chamber of Congress. Borrowing the classification in Barabas et al. (2014), the CES items focus on policy-specific facts whereas the ANES battery tests general institutional knowledge.[11] I combine responses on these items to form an additive index of *factual knowledge* about politics. In order to facilitate easier comparisons with discursive sophistication, each factual knowledge measure is rescaled to zero mean and unit variance. As an additional benchmark, I consider *interviewer assessments* of each respondent's political sophistication (see Bartels 2005; but cf. Ryan 2011).[12]

Figure 1 compares discursive sophistication to conventional knowledge metrics for the CES and 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 ($r = .36$ vs. $r = .23$ in 2016, and $r = .45$ vs. $r = .31$ in 2012), which indicates that the open-ended measure captures characteristics that influence subjective assessments of sophistication. In other words, a respondent's verbatim answers seem to be more influential for subsequent knowledge assessments by

---

[11]See Appendix A for details.

[12]Interviewer assessments were only recorded in the face-to-face sample of the 2012 and 2016 ANES.
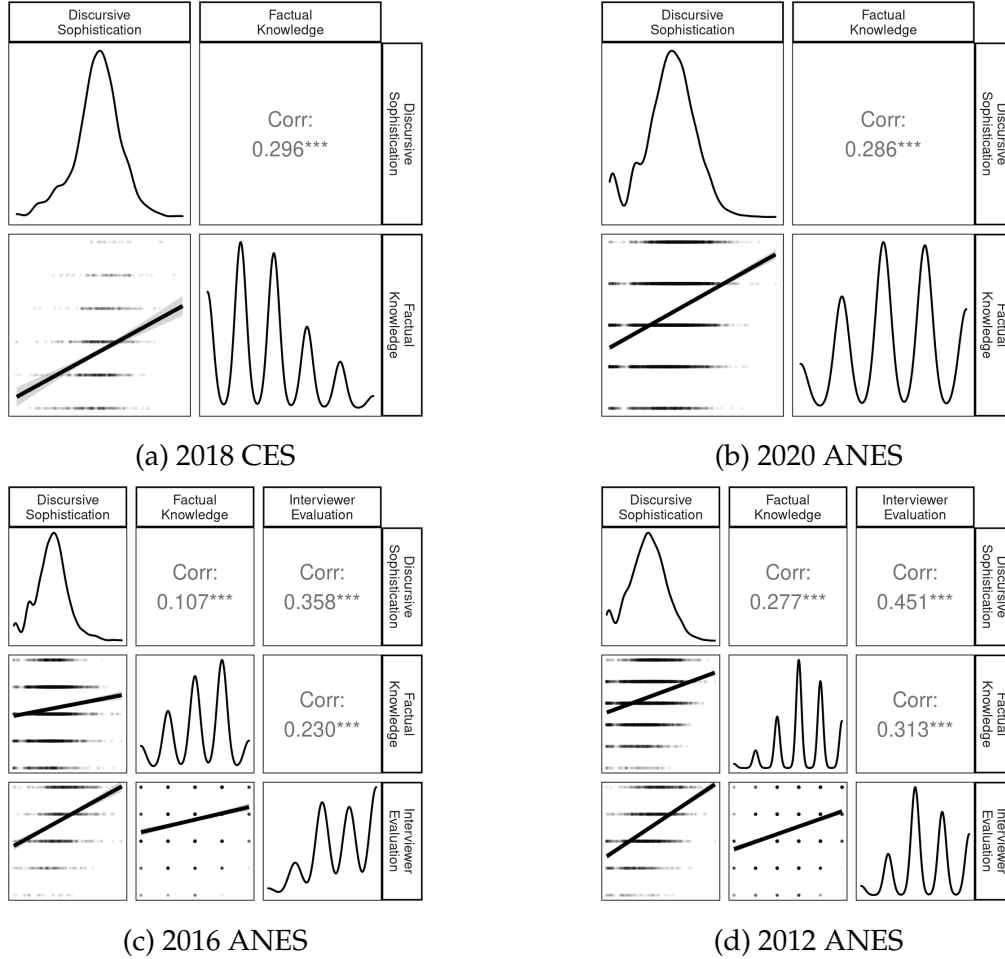
Figure 1: Correlation matrix of discursive sophistication and conventional political knowledge metrics. The plots on the diagonal display univariate densities for each variable. The panels in the lower triangular display the scatter plot of two measures as well as a linear fit. The upper triangular displays the correlation coefficient. All correlations reported are statistically significant with $p < .05$.

the interviewer than a respondent's performance on the factual knowledge questions.

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 from two individuals in the 2018 CES who *scored equally on factual knowledge* (3 out of 5 correct responses), but varied in discursive sophistication.

The results are presented in Table 2. Each row represents one of the open-ended responses targeting specific policy issues. Column A displays the responses of an individ-

	A: Low Sophistication Response	B: High Sophistication Response
Guns (+)	Mental health.	Mental health issues, prior domestic violence conviction.
Guns (-)	None.	Violates second amendment.
Abortion (+)	None.	Right to life, viability of the fetus, ability of the fetus to feel pain, should have been able to access care earlier.
Abortion (-)	Women will have to seek abortions elsewhere which is dangerous.	Women's bodily autonomy and right to decide, comparatively tiny percentage of abortions, usually for health reasons.
Immigration (+)	We help people fleeing violence.	They did not have a choice, many of them have been here since childhood, they do not have another home.
Immigration (-)	Nothing.	Their parents shouldn't have brought them in the first place, whether it's their fault or not, they take jobs from "real" Americans.
Health care (+)	None.	Too expensive, doesn't allow consumers choice in whether or not to purchase healthcare, has increased costs.
Health care (-)	Protect existing conditions.	We need everyone to participate to keep costs down, healthcare is a human right, preexisting conditions should be covered, people should not be bankrupted by healthcare costs.
Trade policy (-)	None.	Protects US jobs and businesses.
Trade policy (+)	It's hurting our economy.	Protects industries that are waning, raises costs for consumers, doesn't actually create jobs or think about the future.
Disc. Soph.	−1.246	0.947

Table 2: Example of open-ended responses for low and high scores on discursive sophistication with equal factual knowledge scores (3 out of 5 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. Note that responses are slightly edited for readability.

ual who scored low on discursive sophistication and column B displays the responses of a high scoring individual. Even though both individuals have the same factual knowledge score, there are systematic differences in their response behavior that suggest disparity in their political sophistication. Overall, respondent A provided a less elaborate response and only focused on a narrow range of issues. Irrespective of whether one agrees with the specific statements, 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

each issue. 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 correlated with political engagement and citizen competence (Lupia, 2006, 2015). Accordingly, I consider how discursive sophistication is associated with (1) engagement and participation in politics, (2) the ability to incorporate new information, and (3) well-justified policy preferences. Appendix C contains robustness checks and supplementary analyses showing, for instance, how discursive sophistication is furthermore predictive of (4) reduced uncertainty about ideological placements of parties and politicians, and (5) higher probabilities to vote based on ideological proximity.

Engagement and Participation in Politics

Any measure of political sophistication should 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 2 compares the effect of discursive sophistication and factual knowledge on four dependent variables related to political engagement: turnout, political interest, 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. In addition to both key predictors, each model controls for gender, education, income, age, race, and church attendance.[13]

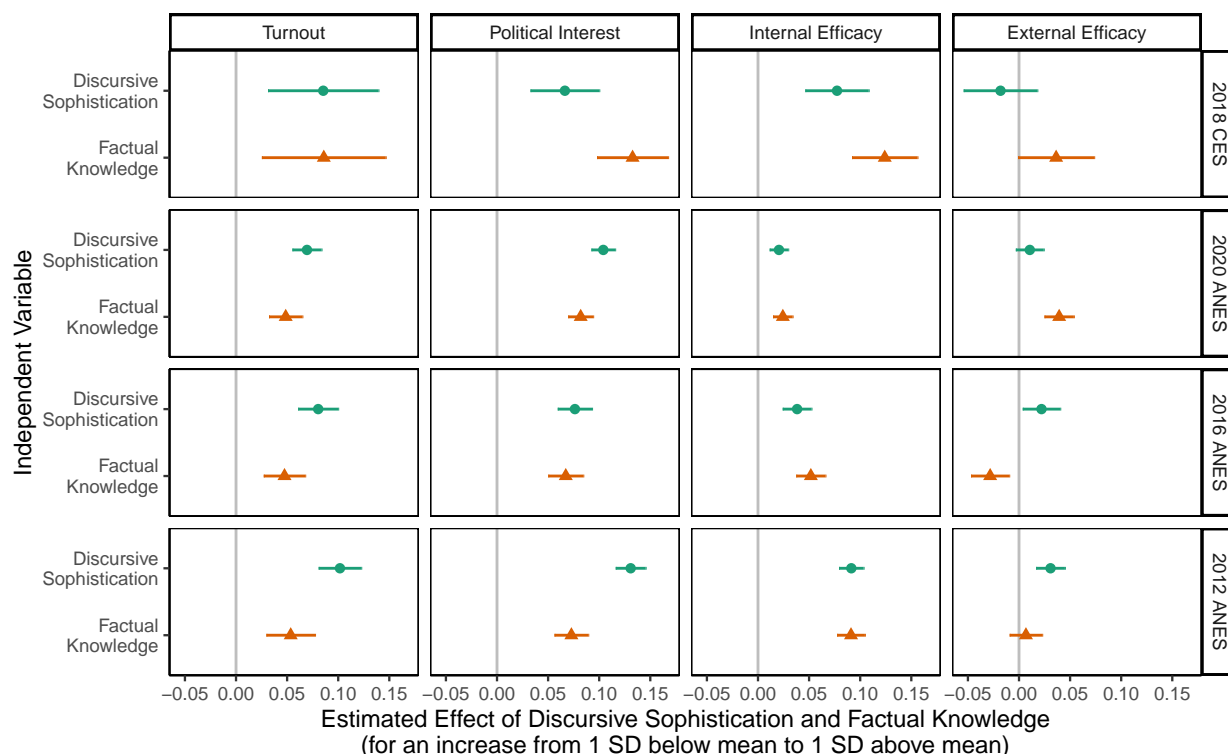[13]See Appendix D for full regression results.

Figure 2: Effects of political sophistication on turnout, political interest, internal efficacy, and external efficacy in the CES and ANES (including 95% confidence intervals). Estimates are based on logistic regression (turnout) or OLS (political interest, internal efficacy, external efficacy). Each model includes controls for sociodemographic variables. Full regression results are displayed in the appendix, Tables D.1 through D.4.

Each panel in Figure 2 compares the estimated effect of increasing either sophistication measure from one standard deviation below the mean to one standard deviation above the mean (holding all other variables constant at their means). Note that the examples previously shown in Table 2 illustrate the substantive meaning of such a two standard deviation increase in discursive sophistication. For factual knowledge, on the other hand, this increase is approximately equivalent to correctly answering three additional knowledge questions.[14] Of course, these effects are purely correlational and should

---

[14]One important difference between both measures is that discursive sophistication is continuous and normally distributed, whereas factual knowledge is only "quasi-continuous" and often skewed (see Figure 1). Regarding the consequences of this difference for the estimates displayed in Figure 2, two related considerations need to be raised. On the one hand, using a truly continuous measure adds variation, which should result in more precise estimates. On the other hand, we have to consider the impact two potential sources of measurement error: 1) estimation uncertainty inherent to discursive sophistication (due to modeling assumptions when processing text as data etc.) and 2) forced discretization inherent to additive

not be interpreted causally. Nevertheless, across all four surveys, discursive sophistication and factual knowledge are complementary and similarly sized predictors of turnout, political interest, and internal efficacy.[15] Only for external efficacy we find more ambiguous results. Factual knowledge has strikingly inconsistent effects—sometimes predicting higher, lower, or no change in external efficacy. Discursive sophistication, in contrast, is more consistently associated with higher external efficacy (the only exception is the 2018 CES, which uses a shorter battery to measure external efficacy).

Considering these initial results, a potential concern may be that discursive sophistication is confounded by individual characteristics that influence verbatim response patterns as well as engagement. As a robustness check, Appendix C.III provides additional regression results controlling for various factors that might drive verbosity such as personality (extraversion, openness to experience, being reserved), survey mode (online vs. face-to-face), verbal skills (Wordsum vocabulary test score), as well as overall verbosity itself (response length). The substantive conclusions remain unchanged.

Incorporation of New Information

In order to replicate and extend this first validation, 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. residents. As part of this study, respondents were asked four open-ended questions to describe their attitudes towards two salient issues: gun legislation and the Affordable Care Act.

knowledge scales (due to measuring a continuous latent construct using a discrete scale). Depending on which of these sources of measurement error is larger, we may see more uncertainty and/or attenuation bias for one metric or the other. Since quantifying and comparing these different sources of measurement error is outside of the scope of this article, I leave this issue for future research.

[15]Additional analyses including interactions between discursive sophistication are included in Appendix D. Interestingly, these models reveal positive and statistically significant main effects for both measures while the interaction coefficients are largely null. There are a few exceptions, however, where we additionally observe *positive* interactions between discursive sophistication and factual knowledge, which suggests that both concepts can be mutually reinforcing. See Table D.1 through D.4 for details. I thank an anonymous reviewer for suggesting these supplementary analyses.

Political sophistication should make it easier for people to incorporate relevant new information about parties, office-holders, and policies. After all, Zaller (1990, 1992) and others argue that factual knowledge is the best available proxy for political awareness. In this analysis, I explore whether discursive sophistication serves as an accurate 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*, ranging from 0 to 9) 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.
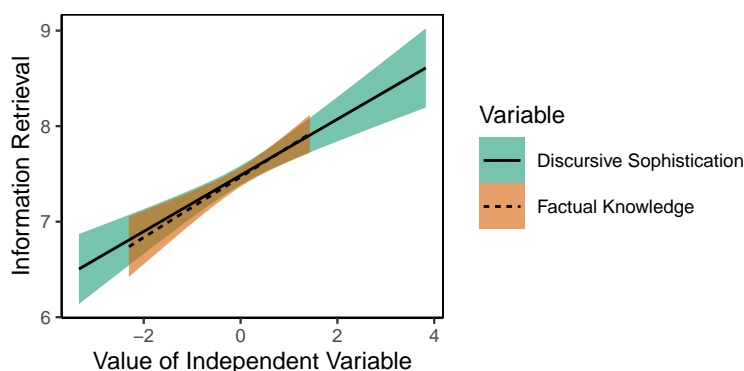

Figure 3: Expected information retrieval in the 2015 YouGov Study as a function of political sophistication (including 95% confidence intervals). Estimates are based on a linear regression including controls for sociodemographic variables. Full regression results are displayed in the appendix, Table D.5.

Figure 3 displays the relationship between political sophistication and disease information retrieval in the 2015 YouGov study. Estimates are based on linear regression 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. Recall that both measures are rescaled to zero mean and unit variance to facilitate direct comparisons be-

tween them. Both discursive sophistication as well as factual knowledge are positively correlated with the amount of information individuals are able to recall from a news article discussing a fictional disease. In addition, this analysis reveals how discursive sophistication can help explain important variation at both tails of the distribution. Conventional additive knowledge scales often suffer from ceiling effects since there is no way to differentiate respondents who answer all questions correctly (or incorrectly, although that is less common with standard batteries). Discursive sophistication suffers from no such constraints and therefore allows us to better represent the full spectrum of the underlying latent variable. Thus, the degree to which citizens discuss their own political beliefs in a more elaborate manner is not only a strong predictor of political engagement but also serves as a powerful proxy for the ability to incorporate new information about a non-partisan issue.

Well-Justified Policy Preferences

As the last validation step, I examine an additional set of surveys that provide a unique opportunity to compare my proposed measure of discursive sophistication with manually coded open-ended responses across three languages. Colombo (2018) compiled a data set of cross-sectional surveys administered in Switzerland after national popular votes on multiple policy propositions. For each referendum, respondents were asked to explain why they voted in favor or against a given proposition in two separate open-ended items. Based on these verbatim responses, I computed discursive sophistication using the same procedure outlined above. Since the survey was conducted in three different languages (German, French, and Italian), I created separate metrics for each group.

Beyond the ability to incorporate new information, political sophistication should enable people to justify their own preferences. Colombo's (2018) manual coding of the respondents' *level of justification* assessed the content, elaboration, and complexity of open-ended responses. Thus, this study provides an opportunity to directly assess the extent to

which high levels of discursive sophistication correspond to well-justified policy preferences in open-ended responses. Any overlap between Colombo's (2018) manual coding with my automated measure corroborates the face validity of discursive sophistication.

The results are presented in Figure 4, which displays the distribution of discursive sophistication for each level of justification coded by Colombo (2018) 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 ($r = 0.23, 0.33$, and 0.36). The proposed measure of discursive sophistication therefore shows a high degree of correspondence with individual levels of justification assessed by independent manual coders.
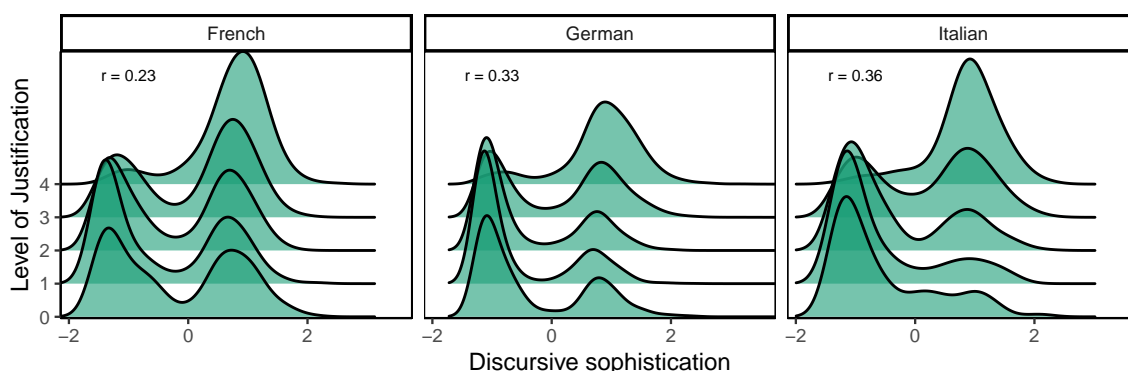

Figure 4: Discursive sophistication and manually coded level of justification (Colombo, 2018) in Swiss post-referendum surveys. The plot compares kernel densities of discursive sophistication for each manually coded level of justification.

To summarize, the results presented thus far indicate that while discursive sophistication shares common characteristics with factual political knowledge measures, both capture different dimensions of sophistication. Indeed, the text-based measure and conventional metrics are independent predictors of political participation and engagement. In addition, discursive sophistication provides a better proxy for the ability to incorporate new information from news sources and shares significant overlap with manually coded levels of justification in open-ended responses. Supplementary analyses reveal that respondents who score higher on discursive sophistication display a smaller degree of un-

certainty around the ideological placement of politicians and parties (Appendix C.IV)—and are ultimately more likely to vote based on ideological proximity in senatorial races (Appendix C.V). Next, I illustrate how discursive sophistication can help refine previous findings regarding the gender gap in 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? Figure 5 displays the distributions of discursive sophistication and conventional metrics comparing both genders. While we observe sizable and statistically significant gender gaps in factual knowledge across the CES, ANES, and YouGov surveys, these differences all but disappear for discursive sophistication. Even though women do not perform as well as men on political quizzes, they do not differ substantially in the level of elaboration when describing their political preferences.

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 (2018) 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 Swiss 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).
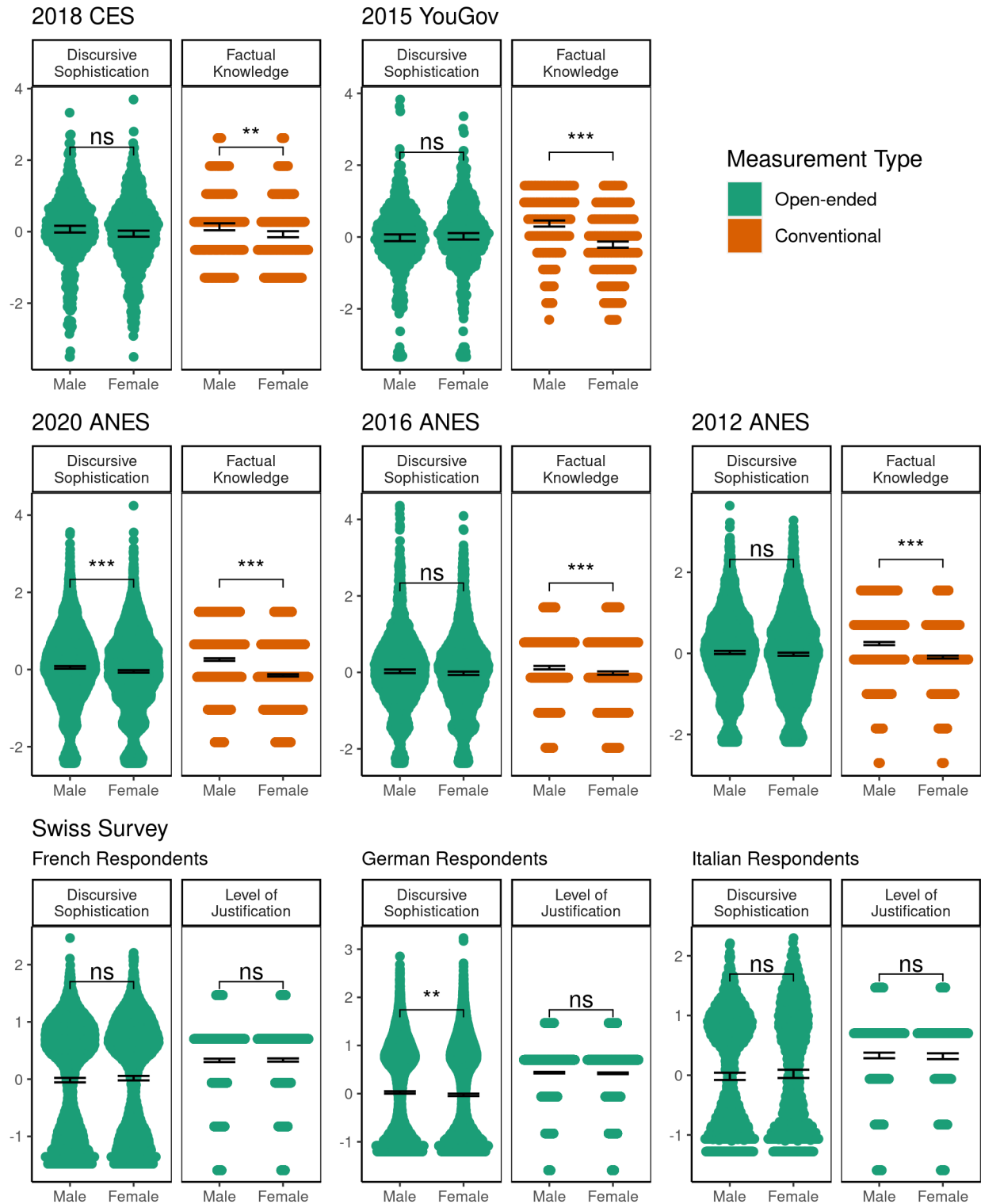
Figure 5: The gender gap in political sophistication. The figures display distributions of political sophistication using open-ended or conventional measures comparing women and men (including 95% confidence intervals around the means). Gender differences are statistically significant at: $^*p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$.

The bottom row of Figure 5 reveals insignificant gender differences for all but one of the metrics across all three languages in the Swiss referendum surveys.[16] Thus, the absence (or at least reduction) of the gender gap remains robust—whether open-ended responses are coded manually or using the discursive sophistication approach.

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. In addition, to the extent that we observe significant gender differences in discursive sophistication—such as in the 2020 ANES or among German respondents in the Swiss survey—we need to assess to what extent these differences are substantively meaningful. Figure 6 shows estimated gender differences after controlling for various potential common determinants such as education, income, age, race and church attendance. Following Rainey (2014), the figure also displays a range of small effect sizes (equivalent to Cohen's $d \leq 0.2$; see Sawilowsky, 2009), in order to evaluate whether *statistically* significant differences are indeed *substantively* meaningful.

After controlling for common determinants, discursive sophistication only reveals negligible (and almost exclusively statistically insignificant) differences between women and men across the CES, ANES and YouGov surveys. Indeed, the 90% confidence intervals only contain negligible effects within Cohen's $d \leq 0.2$, which implies that we can reject the null hypothesis of meaningful gender differences in discursive sophistication (Rainey, 2014).[17] The gender gap in factual political knowledge, however, persists and is substantively as well as statistically significant.[18] Thus, a considerable portion of

---

[16]I will discuss the substantive size of this gender difference in the following section.

[17]The fact that the negligible gender differences observed in the 2020 ANES and among German respondents in the Swiss survey remained statistically significant can be explained by both studies' exceedingly large sample sizes ($N \approx 7,000$ and $N \approx 12,500$, respectively).

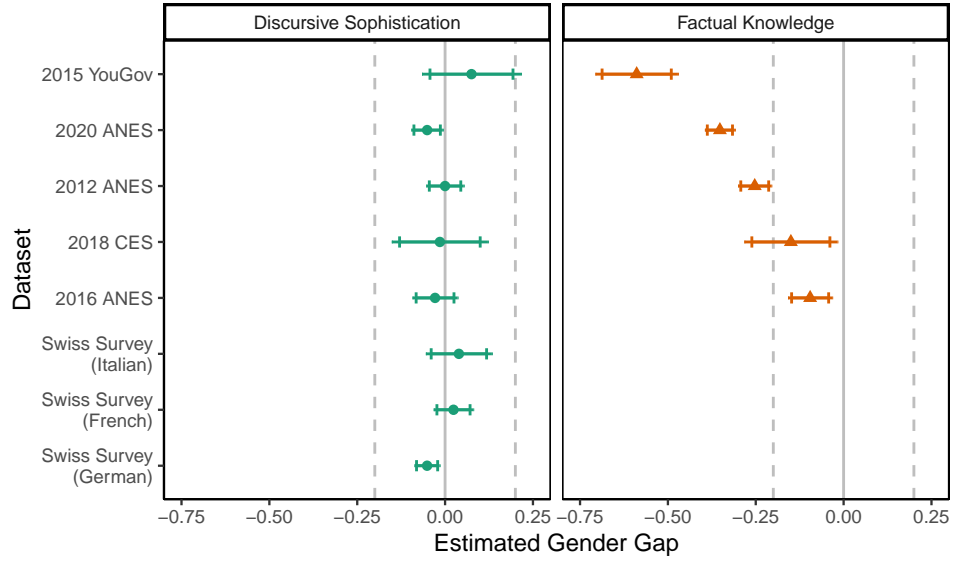[18]Note that the Swiss survey did not include factual knowledge items.

Figure 6: The gender gap in political sophistication controlling for common determinants. Estimates are OLS regression coefficients with 95% and 90% confidence intervals. Dependent variables are discursive sophistication and factual political knowledge. Estimates are based on a linear regression including controls for sociodemographic variables. Dashed lines indicate a range of small effect sizes equivalent to Cohen's $d \leq 0.2$. Full regression results are displayed in the appendix, Tables D.6 and D.7.

the observed differences in factual knowledge between women and men cannot be attributed to underlying disparities in resource-related factors or engagement. Comparing the confidence intervals across both measures further reveals that the insignificant gender differences in discursive sophistication are estimated with similar precision as the significant differences in factual knowledge. Such a result precludes the possibility that null findings for discursive sophistication are purely driven by measurement error on the dependent variable. It is also worth pointing out in this context that the remaining control variables exhibit effects of similar magnitude (and uncertainty) across both measures, which further suggests that there is no systematic difference in measurement error.[19] For instance, knowledge and discursive sophistication are significantly higher among respondents who are more educated and have higher income. The finding that core sociodemographic predictors of political sophistication are consistent across models

---

[19]See Appendix D for full regression results.

lends additional validity to the open-ended measure.

That said, supplementary analyses included in Appendix C.VI reveal that discursive sophistication and factual knowledge have diverging associations with certain personality characteristics, verbal skills, and survey mode.[20] For instance, while openness to experience has a positive effect on discursive sophistication, it has a negative effect on factual knowledge (at least in the 2012 ANES). Being reserved, on the other hand, shows a negative association with discursive sophistication but no relationship with factual knowledge. Especially interesting, however, is the finding that verbal skills (measured using the Wordsum vocabulary test) have a stronger effect on factual knowledge than discursive sophistication. Furthermore, 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 (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. These results illustrate once again that both measures should be seen as complements rather than competing metrics of political sophistication, as they capture different aspects of the underlying concept of interest.

Explaining the (Lack of a) Gender Gap

To summarize, conventional knowledge measures and discursive sophistication produce diverging conclusions regarding the existence of a gender gap. This naturally raises the question which metric we should ultimately trust? Prior research attributed gender differences in factual knowledge—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

---

[20]These analyses are based on the 2012 and 2016 ANES, where additional measures of personality, verbal skills, and survey mode were available.

item 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, which one is more likely to be an artifact of the respective measurement approach: the *existence* of a gender gap in factual knowledge or the *absence* of a gap in discursive sophistication?

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, the surveys presented here 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 presented thus far have shown, this potential selection 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.[21] In fact, adjusting for potential selection effects when examining determinants of sophistication does not change the substantive conclusions.

The second major explanation for the gender gap in political knowledge focuses on question content. By choosing a specific set of recall questions as a general metric for political knowledge, researchers are making strong assumptions about the information

---

[21]See Appendix B for details.

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).

In this last analysis, I therefore explore how women and men differ in topical prevalence across open-ended responses in the 2012, 2016, and 2020 ANES. Note that these open-ended items did not focus on specific issue areas as in the CES, but rather asked respondents to evaluate different political parties and candidates. Thus, they were able to focus on whatever issue they deemed most important. Figure 7 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. The top five topics are more prevalent among men and the bottom five are more likely to be mentioned by women. The label for each coefficient consists of the five highest probability terms related to the topic to illustrate its content.

Taking the 2012 ANES as an example, the topic consisting of terms such as *care*, *health*, and *reform* 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 *tax*, *deficit*, and *cut*. Overall, across all three waves of the ANES, women were less likely than men to

2012 ANES

polici, econom, foreign, platform, weak
record, success, privat, action, sector
tax, deficit, cut, plan, budget
bush, bank, afghanistan, corpor, account
democrat, republican, parti, polit, elect
talk, that, one, lot, thing
live, decis, made, serv, afraid
health, care, insur, reform, medicar
year, done, said, four, chanc
help, rich, poor, tri, peopl

-0.05   -0.04   -0.03   -0.02   -0.01   0.00   0.01   0.02

2016 ANES

polici, foreign, econom, social, affair
money, govern, spend, corpor, self
court, secur, suprem, dishonesti, respons
honesti, crimin, leadership, behavior, integr
christian, public, abil, base, connect
femal, negat, littl, seem, theyv
just, say, that, guess, thing
presid, woman, first, america, great
minor, women, ignor, comment, toward
peopl, countri, tri, help, want

-0.05   -0.04   -0.03   -0.02   -0.01   0.00   0.01   0.02

2020 ANES

border, trade, economi, china, illeg
court, suprem, use, solv, solut
balanc, special, strength, interest, voic
democrat, left, radic, parti, far
truth, leadership, honesti, integr, tell
black, thought, matter, communiti, live
stand, everyth, god, love, freedom
class, rich, middl, help, peopl
disrespect, women, attitud, bulli, honest
think, dint, like, say, lot

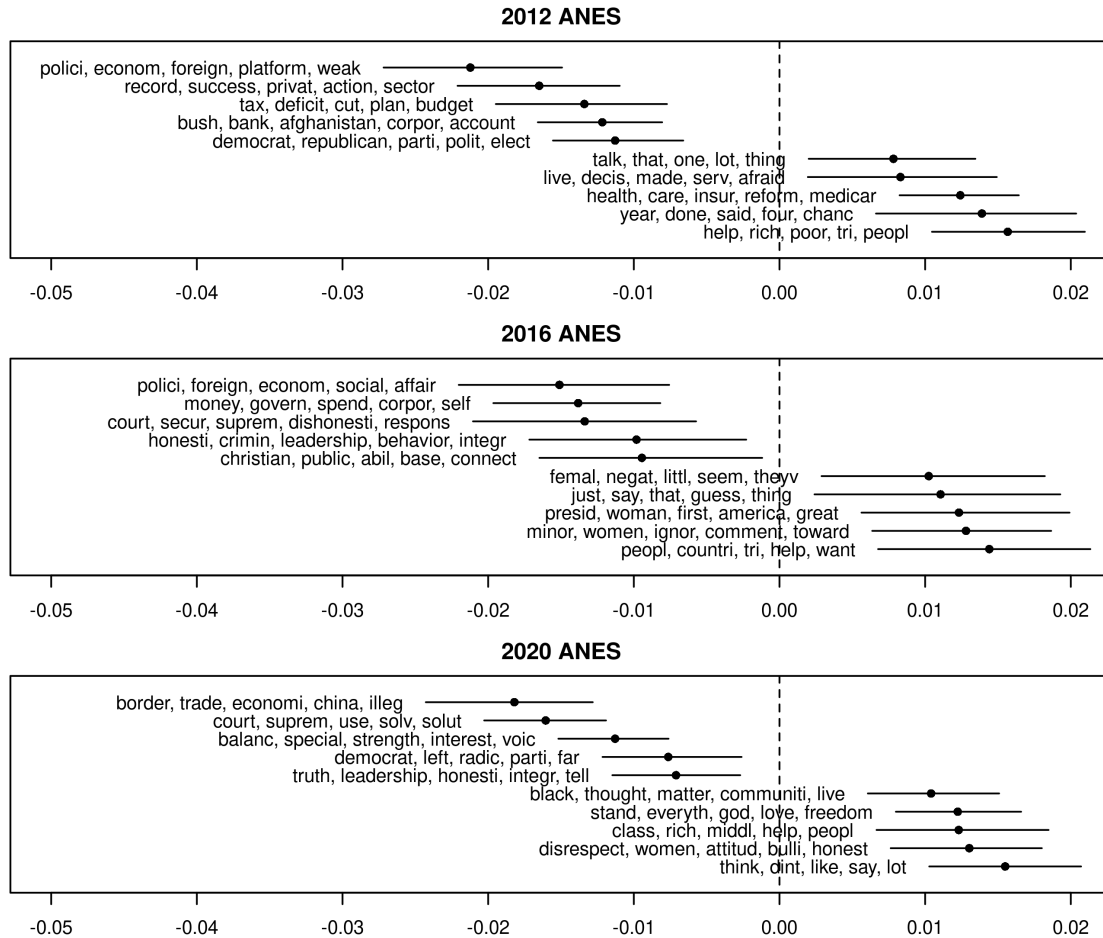-0.05   -0.04   -0.03   -0.02   -0.01   0.00   0.01   0.02

Figure 7: 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 most frequent and exclusive (FREX) terms associated with each topic.

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 govern-

31

ment 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 are at least partly driven 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 disappears since open-ended questions about political preferences allow respondents to focus on specific considerations that are most salient to them.

Discussion

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 political competence. 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 appropriately in their decision-making process. 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; Ferrín et al., 2020), but it cannot be properly addressed while relying exclusively on off-the-shelf recall questions to measure political knowledge. What is more, our discipline lacks a 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 provide a first step towards devising balanced recall items. More specifically, examining the types of issues women and men emphasize when discussing their political preferences in open-ended responses can serve as a guide to select new closed-ended item batteries. Building on this argument, Kraft and Dolan (2022a) show that focusing on issues emphasized by both women and men all but diminishes the gender gap in factual knowledge. Thus, researchers could use this heuristic—either by relying on publicly available surveys containing open-ended items, or by fielding them in a pilot study—in order to select gender-balanced knowledge questions for their survey.

Of course, relying on open-ended responses to assess political sophistication has its limitations, too. First and foremost, elaboration in verbatim attitude expression may be 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 potential confounding factor that is unique to open-ended responses is the respondents' general linguistic skills or verbal verbosity, which may 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 to the inclusion of measures of linguistic skills or personality characteristics such as extraversion. In a similar vein, the gender gap finding did not appear to be driven by selection effects, which again sug-

gests 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 Appendix C.VI). 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.

A closely related concern pertains to the question whether discursive sophistication captures uniquely *political* skills or can rather be viewed as a more general phenomenon. In other words, are we simply measuring basic communication skills rather that transcend the realm of politics? To provide an initial answer to this important question, recall that the supplementary analyses in Appendix C show that 1) core results reported in the article hold after controlling for verbal skills and general verbosity, and 2) discursive sophistication is predictive of distinctly political competences such as the degree of certainty around the ideological placement of politicians and parties or people's likelihood to vote based on ideological proximity in senatorial races. On the other hand, recent research suggests that factual knowledge about politics is itself not domain-specific and largely resides on the same dimension as other knowledge topics such as sports or popular culture (Burnett and McCubbins, 2018). Nevertheless, future studies should further assess this question empirically—for instance by comparing discursive sophistication based

on open-ended responses centered around politics with an equivalent measure based on non-political questions such as sports, literature, or science. Going forward, this line of research could develop best practices for the selection of different question types (e.g., targeting specific policies vs. party evaluations) to measure discursive sophistication.

Lastly, a skeptic may still argue that while open-ended responses may provide useful insights, manual coding is still 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 (2018) 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.

Conclusion

Political scientists should worry less about pure levels of *factual knowledge* and instead focus on how people justify their political preferences. 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 or issues. 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. 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. Furthermore, a free soft-

ware package for the statistical programming environment R allows applied researchers to implement the framework in their own surveys.[22]

The findings presented in this article 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 and factual knowledge are independent predictors of political engagement and efficacy. The text-based measure is furthermore 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. This lack of gender differences in discursive sophistication can be attributed to the fact that open-ended questions allow women to focus on different considerations than men.

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 article 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.

---

[22]R package available here: ███████████████████████████████████.

References

Abrajano, Marisa. 2014. "Reexamining the "Racial Gap" in Political Knowledge." *The Journal of Politics* 77 (1): 44–54.

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.

Bernhard, Rachel, and Sean Freeder. 2020. "The more you know: Voter heuristics and the information search." *Political Behavior* 42 (2): 603–623.

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.

Bullock, John G., and Kelly Rader. 2021. "Response Options and the Measurement of Political Knowledge." *British Journal of Political Science*: 1–10.

Burnett, Craig M, and Mathew D McCubbins. 2018. "Is political knowledge unique?" *Political Science Research and Methods* 8 (1): 188–195.

Clifford, Scott, and Jennifer Jerit. 2016. "Cheating on Political Knowledge Questions in Online Surveys: An Assessment of the Problem and Solutions." *Public Opinion Quarterly* 80 (4): 858–887.

Colombo, Céline. 2018. "Justifications and Citizen Competence in Direct Democracy: A Multilevel Analysis." *British Journal of Political Science* 48 (3): 787–806.

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.

Dancey, Logan, and Geoffrey Sheagley. 2013. "Heuristics behaving badly: Party cues and voter knowledge." *American Journal of Political Science* 57 (2): 312–325.

Dassonneville, Ruth, Mary Nugent, Marc Hooghe, and Richard R Lau. 2020. "Do Women Vote Less Correctly? The Effect of Gender on Ideological Proximity Voting and Correct Voting." *Journal of Politics* 82 (3): 1156–1160.

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. 1993. "Measuring political knowledge: Putting first things first." *American Journal of Political Science* 37 (4): 1179–1206.

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.

Ferrín, Mónica, Marta Fraile, and Gema García-Albacete. 2017. "The Gender Gap in Political Knowledge: Is It All about Guessing? An Experimental Approach." *International Journal of Public Opinion Research* 29 (1): 111–132.

Ferrín, Monica, Marta Fraile, and Gema García-Albacete. 2018. "Is It Simply Gender? Content, Format, and Time in Political Knowledge Measures." *Politics & Gender* 14 (2): 162–185.

Ferrín, Monica, Marta Fraile, Gema M García-Albacete, and Raul Gómez. 2020. "The gender gap in political interest revisited." *International Political Science Review* 41 (4): 473–489.

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. 2014a. "Do women know less about politics than men? The gender gap in political knowledge in Europe." *Social Politics: International Studies in Gender, State & Society* 21 (2): 261–289.

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

Fraile, Marta, and Jessica Fortin-Rittberger. 2020. "Unpacking gender, age, and education knowledge inequalities: A systematic comparison." *Social Science Quarterly* 101 (4): 1653–1669.

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.

Höhne, Jan Karem, Carina Cornesse, Stephan Schlosser, Mick P Couper, and Annelies G Blom. 2020. "Looking up answers to political knowledge questions in web surveys." *Public Opinion Quarterly* 84 (4): 986–999.

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

Kraft, Patrick W. 2018. "Measuring Morality in Political Attitude Expression." *The Journal of Politics* 80 (3): 1028–1033.

Kraft, Patrick W, and Kathleen Dolan. 2022a. "Asking the Right Questions: A Framework for Developing Gender-Balanced Political Knowledge Batteries." *Political Research Quarterly*: 10659129221092473.

Kraft, Patrick W, and Kathleen Dolan. 2022b. "Glass Half Full or Half Empty: Does Optimism about Women's Representation in Elected Office Matter?" *Journal of Women, Politics & Policy*: 1–13.

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.

Lau, Richard R, David J Andersen, and David P Redlawsk. 2008. "An exploration of correct voting in recent US presidential elections." *American Journal of Political Science* 52 (2): 395–411.

Lau, Richard R., and David P. Redlawsk. 2001. "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45 (4): 951-971.

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

Lodge, Milton, and Charles S Taber. 2013. *The Rationalizing Voter*. Cambridge University Press.

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.

Luskin, Robert C. 1990. "Explaining political sophistication." *Political Behavior* 12 (4): 331–361.

Luskin, Robert C, and John G Bullock. 2011. ""Don't know" means "don't know": DK responses and the public's level of political knowledge." *The Journal of Politics* 73 (02): 547–557.

Macdonald, Stuart Elaine, George Rabinowitz, and Ola Listhaug. 1995. "Political Sophistication and Models of Issue Voting." *British Journal of Political Science* 25 (4): 453-483.

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

McAllister, Ian. 2019. "The gender gap in political knowledge revisited: Australia's Julia Gillard as a natural experiment." *European Journal of Politics and Gender* 2 (2): 197–220.

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.

Pennebaker, James W, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. Technical report.

Pereira, Frederico Batista. 2019. "Gendered Political Contexts: The Gender Gap in Political Knowledge." *The Journal of Politics* 81 (4): 1480–1493.

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.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58 (4): 1083–1091.

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.

Sawilowsky, Shlomo S. 2009. "New effect size rules of thumb." *Journal of modern applied statistical methods* 8 (2): 26.

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.

Style, Hillary, and Jennifer Jerit. 2020. "Does it matter if respondents look up answers to political knowledge questions?" *Public Opinion Quarterly* 84 (3): 760–775.

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

Tausczik, Yla R, and James W Pennebaker. 2010. "The psychological meaning of words: LIWC and computerized text analysis methods." *Journal of language and social psychology* 29 (1): 24–54.

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.

Vegetti, Federico, and Moreno Mancosu. 2020. "The Impact of Political Sophistication and Motivated Reasoning on Misinformation." *Political Communication* 37 (5): 678–695.

Verba, Sidney, Nancy Burns, and Kay Lehman Schlozman. 1997. "Knowing and caring about politics: Gender and political engagement." *The Journal of Politics* 59 (4): 1051–1072.

Wolak, Jennifer. 2020. "Self-Confidence and gender gaps in political interest, attention, and efficacy." *The Journal of Politics* 82 (4): 1490–1501.

Wolak, Jennifer, and Eric Gonzalez Juenke. 2021. "Descriptive representation and political knowledge." *Politics, Groups, and Identities* 9 (1): 129–150.

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.