Political Analysis (), , 1–37

ARTICLE

Text as Behavior

Omar Wasow*

Department of Political Science, UC Berkeley
*Corresponding author. Email: owasow@berkeley.edu

Abstract

Text analysis typically focuses on content—such as sentiment or topic—but expression is also a form of effortful action. Building on this insight, I propose using simple features of open-ended tasks to study *text as behavior.* This approach treats expression, such as writing, as cognitively, emotionally and temporally "costly" for subjects but inexpensive for researchers. I show basic statistics like the number of characters can approximate effort and significantly improve estimation of quantities of interest, including candidate choice, the probability of turning out to vote, and psychological states about which a subject may not be fully aware. Further, these methods can convert nonresponse into informative data; validate survey instruments; serve as mechanism checks; be hard for a subject to "game"; work across different languages; and analogize well to real-world situations. In sum, text as behavior can help address a range of issues related to quantifying attitudes and actions.

Keywords: Text analysis, text as data, political behavior, survey methods

Word count: 5,882

1. Introduction

Writing is hard. Linguist Peter Hugoe Matthews (2003, 16) said of humanity, "No one would describe us as the 'writing' species." At the same time, writing is a pervasive part of contemporary life. The rising ubiquity of writing follows long-term educational and technological trends. In 1900, only one in five people in the world was literate. By the end of the twentieth century, only one in five people was *illiterate* (van Zanden et al. 2014). Further, as Thompson (2010) notes, "This generation doesn't make phone calls, because everyone is in constant, lightweight contact in so many other ways: texting, chatting, and social-network messaging." In contrast to Matthews' statement two decades ago, now, for many modern, connected and educated people, "Homo scribens" might actually be an appropriate description.

Quantifying sincere attitudes is also hard and a central challenge for social science (Campbell and Stanley 1963). Researchers often attempt to gauge affect and behavior through games or by attaching costs or rewards to an action in an effort to encourage subjects to reveal genuine tendencies and tastes (McDermott 2002). These approaches, though, have limitations. Paying subjects additional money to reveal preferences can be expensive. Economic games, such as the "Dictator Game," may not translate well to real-world situations in which subjects are not participants in a study (Winking and Mizer 2013). In addition, many measures suffer from concerns about external validity (Findley, Kikuta, and Denly 2021).

Given the cognitive, affective and temporal demands of expression, I propose using simple metadata from open-ended prompts—like nonresponse and number of characters—as alternative measures of effortful action. The nearly universal difficulty of writing combined with its increasing prevalence allows for a measure that is "costly" for subjects but inexpensive for researchers. As Berinsky (2013) noted, "respondents must pay costs—albeit small—to form and express their views in a survey" (9). Treating expression as a behavioral measure extends Berinsky's insight to open-ended text. Open-ended prompts can also be evaluated in ways that are unlikely to be anticipated by subjects and, in many cases, these tasks correspond neatly to many real-world behaviors. I test this *text as behavior* method across a range of topics, datasets, and languages with three simple features of open-ended prompts: character counts, nonresponse and time.

In contrast to methods like sentiment analysis that attempt to extract meaning from terms, I use metadata to determine often-unobserved qualities like effort, preference intensity and ambivalence. I find these simple features of text provide meaningful signals of attitudes and behaviors. All other factors being equal, subjects who are more expressive reveal significantly more intense support, engagement and capacity for relevant action related to that particular topic as compared with subjects who communicate less. Cavaillé, Chen, and Straeten (2024), for example, tested whether conventional survey scales might collapse important variation in how much people care about particular issues. They applied the text as behavior approach and found that character counts in letter-writing tasks on policies like the minimum wage and abortion served as an effective behavioral measure of preference intensity.

Further, I find that subjects with blank responses—whether from skipping a question or replying "No" when prompted to elaborate—tend to exhibit more negative attitudes and conflicted behaviors compared to subjects who express even a single word. While the motivating idea stems from the cognitive demands of writing, I test whether expressive behavior more broadly—typed or transcribed from speech—provides useful signals. In practice, I find that both written and spoken responses yield meaningful results, suggesting the underlying dynamics are not limited to any one mode of expression.

This approach is both a complement to and offers some advantages over many current and more sophisticated methods of measuring attitudes and behavior. First, the use of open-ended prompts is already widespread and growing in social science (Li 2023). Consequently, applying text as behavior methods to existing or planned open-ended prompts adds minimal cost or respondent burden to data collection, yet captures some of the cognitive, affective, and temporal demands faced by subjects. In contrast, while adding more survey items can increase reliability, doing so typically raises costs and respondent fatigue.

Second, text responses—whether collected in surveys, transcripts or social media—can offer researchers good equivalence to real-world situations or "mundane realism" (Aronson and Carlsmith 1968). Third, the demands of expression allow measurement of otherwise hard-to-quantify traits such as the degree of commitment to expressing oneself on political issues and in voting. Fourth, expression can offer a window into psychological states about which a subject may not be fully aware (Wilson 2004). Fifth, blank answers may be informative rather than missing data.

Sixth, the open-ended nature of text responses—and the nearly infinite number of ways they can be quantified—make them difficult to "game." Rather than relying on what respondents say, metadata about *how* they respond, such as length or latency, acts like a behavioral "derivative," capturing traces of engagement apart from content. Because respondents are typically unaware these features are being measured, such implicit indicators benefit from what we might call *metric opacity*, making them less susceptible to social desirability or other forms of strategic responding.

Seventh, any measure of human attitudes or behavior risks participants inferring the study's purpose and altering their behavior in response (Mummolo and Peterson 2019). Metadata from open-ended prompts, by contrast, provide *hypothesis opacity*, reducing the likelihood that respondents anticipate or adapt to researcher expectations. Finally, text as behavior appears to generalize across languages, modes, transcription, and translation.

2. Related Work

Social scientists now routinely study text as data (Grimmer, Roberts, and Stewart 2022; Li 2023) using methods such as topic models (Roberts et al. 2014) and forms of sentiment analysis (cf. Mossholder et al. 1995; Bisgaard 2019). In contrast, treating text as a behavioral measure remains less common, though it builds on at least five related bodies of scholarship.

First, scholars employ text to measure emotions, mental states, how subjects acquire information and attributes like political sophistication (Pennebaker, Francis, and Booth 2001; Rude, Gortner, and Pennebaker 2004; Kramer, Guillory, and Hancock 2014; Gillion 2016; Benoit, Munger, and Spirling 2019; Bernhard and Freeder 2020; Kraft 2023). Kuo, Malhotra, and Mo (2017), henceforth KMM, experimentally induced feelings of exclusion and, among other measures, asked subjects to write lists of items they liked and disliked about both the Democratic and Republican Parties. Of particular relevance, KMM note, "This task required a great deal of effort on the part of respondents and therefore can be interpreted as a behavioral manifestation of liking or aversion toward a political party" (27). KMM also found that treated Asian American subjects took significantly longer to complete the survey than expected from additive effects alone (see Section A4).

Second, researchers analyze text from transcripts, court cases, social media, short messaging services, and Internet searches as indicators of bias that may circumvent efforts to offer socially acceptable answers. Scholars draw on Google queries, text logs or surveilled and transcribed speech to identify unvarnished, non-survey alternative indicators of attitudes (Stephens-Davidowitz 2014; Maloney 2021). Email- and messaging-based correspondence studies offer an additional way to detect bias solely through written responses (Butler and Broockman 2011; Lowande and Proctor 2020; Yan and Bernhard 2023). Scientists also study deliberative discussions, Supreme Court interruptions and police stops to understand gender and racial bias (Karpowitz, Mendelberg, and Shaker 2012; Jacobi and Schweers 2017; Voigt et al. 2017).

Third, many studies rely on social media to measure attitudes of the mass public and predict a range of future behaviors from voting in elections to movie attendance and disruptive events (cf. Tumasjan et al. 2010; Alsaedi, Burnap, and Rana 2017; Eady, Hjorth, and Dinesen 2022; Wasow et al. 2010). Researchers also observe interactions that reveal debate strategies, dynamics of status hierarchies, polite or conflictual conversations, and perceived violence (Danescu-Niculescu-Mizil et al. 2013; Panteli 2002; Jann and Schottmüller 2024). This line of work also includes experiments in which both treatments and outcomes are short written exchanges (Munger 2017; Mosleh et al. 2021).

Fourth, psychologists routinely employ writing tasks as a form of treatment to induce different states of mind, from reducing anxiety and trauma to increasing awareness of certain moral frames (Pennebaker 1997; Day et al. 2014). Open-ended responses also validate and extend survey instruments (ten Kleij and Musters 2003), serve as a type of mechanism check (KMM), and detect inattentive subjects (Ziegler 2022). Survey methodologists also analyze character-length metadata to identify mode effects between paper and online surveys (Denscombe 2008).

While these four literatures treat text as a behavioral signal, the studies are more substantive than methodological. As a result, they offer limited guidance for applying the approach. More broadly, this scholarship lacks a generalizable model of the cognitive, emotional and temporal demands of expression that could help unify seemingly disparate behaviors, from nonresponse to verbose writing.

Finally, research on survey nonresponse and missing data is also relevant to understanding situations in which subjects opt out of responding or convey nothing (Berinsky 2008, 2013). Scholarship on nonresponse, though, often treats the topic as a form of missing data rather than as a distinct and meaningful form of response. Longford (2007), for example, writes:

> "Methods for addressing nonresponse can be divided into two categories: those that reduce the dataset (by deleting the records of some units) and those that make up the data so as to generate, structurally, a look-alike of the complete dataset" (380).

Longford's two categories, however, exclude a third possibility: *informative nonresponse*.[1] As I show later, zero-character responses are often *not* missing data in the traditional sense. Rather, they convey meaningful information that would be lost or destroyed through standard methods like listwise deletion or imputation.

3. Text as Behavior

I propose that much of this otherwise disparate scholarship can be unified under the category of *text as behavior.* Text as behavior is a subset of text as data, focusing on cases where writing or transcribed speech reflects a modestly taxing action (Berinsky 2013) and, consequently, can help to reveal, validate or predict attitudes, preferences and behaviors. Three bodies of scholarship provide the theoretical foundation for the text as behavior approach: (1) expression is often cognitively, affectively and temporally demanding; (2) expressive behavior often follows systematic patterns that reflect latent orientations—such as ideology or group identities—and can function as a behavioral signature; (3) expressive tasks can provide a partial window into inaccessible, nonnormative or difficult-to-articulate thoughts and feelings.

On the cognitive and affective demands of writing, Hayes (1996) details the complex range of capacities that must be coordinated and executed, from visual and motor skills to short-term memory and language abilities. Kellogg (1999) argues that writing does "not simply unfold automatically and effortlessly in the manner of a well learned motor skill . . . writing anything but the most routine and brief pieces is the mental equivalent of digging ditches" (17, quoted in Graham 2018). Further, expression is often not just cognitively challenging but emotionally hard, too. Consider, for example, sympathy cards with pre-written inscriptions to help solve the problem of conveying an emotion when one feels at a loss for words. Reviewing several decades of psychology research on the relationship between feelings and thoughts, Wright (2017) concludes there is a "fine entanglement of affect and cognition" (120). Lastly, all forms of expression, whether written or spoken, impose a cost in time.

Metadata from interactions can also be revealing. As early as 1984, Michael Crichton's short story *Mousetrap* suggested that behavioral traces in typing and mouse use might be as distinctive as fingerprints (Crichton 1984). While fictional, the narrative prefigures real-world work on what is now called behavioral biometrics. This field uses "the way people do things such as speaking (voice), writing (signature), [and] typing (keystroke dynamics)" to verify identity, often in "stealth mode" (Teh, Teoh, and Yue 2013, 2).

In social science, researchers have adapted similar behavioral traces. Response latencies underpin Implicit Association Tests; mouse-tracking captures moment-to-moment hesitation, conflict or confusion (Horwitz, Kreuter, and Conrad 2017); and eye-tracking shows how people allocate attention when evaluating candidates (Jenke et al. 2021). While these methods may not provide an individual-level biometric "fingerprint," they can yield group-level behavioral signatures, such as "leaning Republican," even when self-reported party identification differs.

Expressive tasks may also provide insight into the nonconscious perceptual systems used to make sense of and interpret the world (Wilson 2004). Haidt (2012) offers a useful metaphor, suggesting the mind is divided like "a rider on an elephant." The rider is that about which we are aware, our conscious reasoning, while the elephant is "the other 99 percent of mental processes—the ones that occur outside of awareness but that actually govern most of our behavior" (xxi). Writing techniques like journaling, free association, and automatic writing have all been suggested as tools for surfacing nonconscious thoughts. Though researchers have developed many creative instruments and games

1. The term "informative nonresponse" is sometimes used in related work typically to describe systematic missingness requiring techniques like imputation (cf. Laaksonen and Chambers 2006). I use the term differently here, to refer to blank responses that nevertheless convey meaning.

to reveal otherwise subterranean thoughts and feelings, open-ended prompts remain underutilized in social science as a window into nonconscious thought processes (Roberts et al. 2014).

4. Data and Methods

Table 1 summarizes three studies demonstrating diverse but related applications of text as behavior. Studies 1–2 draw on the 2016 ANES base sample; for respondents recontacted, I also use their 2020 and 2024 responses. Study 3 uses the 2016 Afrobarometer.

Table 1. Overview of Studies, Outcomes, Predictors, and Generalizable Applications.

Study	Outcomes	Predictor	Generalizable Applications
1a	**Candidate Choice** (2016, 2020, 2024; ANES)	Informative Nonresponse 2016	*Behavioral affect:* Blank *for/against* responses reflects approval intensity toward candidates and parties.
1b	**Candidate Choice** (2016, 2020, 2024; ANES)	Expressive Alignment 2016	*Behavioral affect*: Verbosity in *for/against* prompts reflects the intensity of approval or disapproval toward candidates and parties.
2	**Validated Vote** (2016, 2020; ANES)	Expressive Engagement 2016	*Articulated concern as civic behavior*: Verbosity about most important problems signals capacity for political action.
3	**Democracy Prompt Length** (2016; Afrobarometer): # Characters: "What does democracy mean to you?"	Importance of Democracy Scale 2016	*Multilingual instrument validation*: Verbosity about democracy confirms survey battery about democracy across languages and translation.

4.1 American National Election Study

The ANES surveyed a cross-section of eligible United States voters both before and after the 2016, 2020 and 2024 elections. The 2016 survey was conducted with both a face-to-face sample (N=1,180) and an Internet sample (N=3,090). In addition to a large battery of multiple-choice survey questions, it included some open-ended prompts. Of particular interest were four items that began with yes/no questions such as, "Is there anything in particular about Hillary Clinton that might make you want to vote against her?" Respondents who answered "Yes" were prompted to elaborate in an open-ended response: "What is that?"

In addition, the 2016 ANES included three open-ended "Most Important Problem" prompts asking respondents to identify and prioritize issues facing the country. A fourth question asked, "Which among mentions is the most important problem?" Also, the 2016 and 2020 ANES provided supplemental data with validated turnout incorporating match probabilities using publicly available voter files (Enamorado, Fifield, and Imai 2017). Lastly, a subset of 2016 subjects were recontacted for the 2020 ANES (N $\approx$ 2,839), and 2024 ANES (N $\approx$ 2,171). I use the panel structure to test whether 2016 measures can also predict outcomes in 2020 and 2024.

4.2 Afrobarometer

The 2016 Afrobarometer (Round 6) is a pan-African, non-partisan survey on democracy, governance, and society conducted in 36 countries. Interviews with 53,921 respondents were carried out in native languages and translated into English (34,838 cases), French (14,116), or Portuguese (4,693). Respondents answered three open-ended prompts—"What, if anything, does 'democracy' mean to you?"—and a battery of closed-ended items on the importance of democratic governance. Study 3 uses these data to test whether simple text-metadata measures (e.g., character counts) can validate survey instruments across languages, transcription and translation.

4.3 Defining Measures

Across all studies, I use one or more of three measures: the number of characters in an open-ended response, blank responses (zero characters), and/or completion time for the open-ended task. The specific scales are explained in more detail within each study and follow a common five-step order of operations, summarized below. Steps 2–3 apply only to continuous text measures and are skipped for binary indicators and time (see Section A4).

1. *Count*: For each open-ended response, compute the raw number of characters.
2. *Deskew*: Apply the inverse hyperbolic sine (IHS) transformation to reduce skew and accommodate zeros.
3. *Normalize*: Divide the transformed values by the maximum observed within each survey mode to address mode-specific ceilings (e.g., face-to-face vs. web; see Section A5.1).
4. *Pool*: Combine responses from related prompts. Use addition for prompts with common valence (e.g., "for Clinton" and "against Trump") or unidirectional constructs (e.g., overall engagement) and subtraction for directional constructs (e.g., difference in candidate affect).
5. *Conceptualize*: Use techniques like correlation analysis or principal component analysis (PCA) to better understand and define the construct or scale.

To simplify notation across studies, I define a transformation function $T\colon \mathcal{S} \to \mathbb{R}$, where $\mathcal{S}$ is the set of open-ended responses. This function maps a text string s to a normalized, transformed score in three steps: (1) compute the number of characters in s, denoted nchar(s); (2) reduce skew using IHS function, asinh($\cdot$); and (3) normalize by dividing by the maximum transformed value observed within different survey modes.[2]

$$T(s) = \frac{\operatorname{asinh}(\operatorname{nchar}(s))}{\max_{\text{mode}} \operatorname{asinh}(\operatorname{nchar}(s))} \tag{1}$$

After transformation, each response takes on a value between 0 and 1. Like a log, the IHS compresses large values more than small ones and reduces skew. I opt for the IHS over the log, however, because small values are not compressed as severely. This better distinguishes blank from short responses and the IHS handles zeros gracefully, which is helpful for character counts. These transformations harmonize responses across modes with distinct character limits—for example, 60 online versus 1,000 face-to-face in the 2016 ANES—making it possible to combine measures arising from distinct data-generating processes and extract a meaningful signal. Diagnostic tests indicate that residual differences are small (see Section A5.1).

I opt for the number of characters rather than number of words or stemmed terms on the assumption that expression requires effort and, therefore, each keystroke or character is the most granular measure of exertion expended by a subject. Further, other plausible measures, such as counting terms, necessarily discard information when character lengths differ across terms (e.g., "jobs" versus "unemployment"). Finally, the number of characters succinctly captures the difference between blank responses (zero characters) and response (at least one character). Other reasonable modeling choices and processing orders may be appropriate. I discuss some of these considerations in more detail in Section A5.

2. In the 2016 ANES, the maximum number of characters before transformation was 1,024 for face-to-face and 60 for web respondents. The IHS function and normalization substantially reduce this difference and results are robust other reasonable modeling choices.

5. Study 1: Text as Predictor of Candidate Choice

Can text metadata, independent of content, predict outcomes like vote choice or party identification? In Study 1, I test whether nonresponse and the number of characters in four open-ended candidate evaluation prompts can improve prediction of key political outcomes.

5.0.1 Study 1a: Informative Nonresponse on Candidate Choice

In the first test, I investigate whether declining to respond can still convey meaning. For writing or speaking tasks, nonresponse could be informative, such as with "ghosting" via text or phone. Discussing narrative in political science, Patterson and Monroe (1998) note: "Silences and gaps can be as telling as what is included. ... Like Sherlock Holmes's silent dog—which did not bark because it knew the intruder—the absence of comment may speak volumes. The challenge for the analyst is to interpret what this silence signifies" (329).

In this case, respondents were first asked whether anything would make them want to vote for or against each candidate; only those who answered "Yes" were prompted to elaborate. Both respondents who said "No" and those who said "Yes" but wrote nothing record zero characters in the open-ended response field. Though these differ procedurally, they yield equivalent data. I therefore treat an explicit "No" and a blank response as forms of *informative nonresponse*.

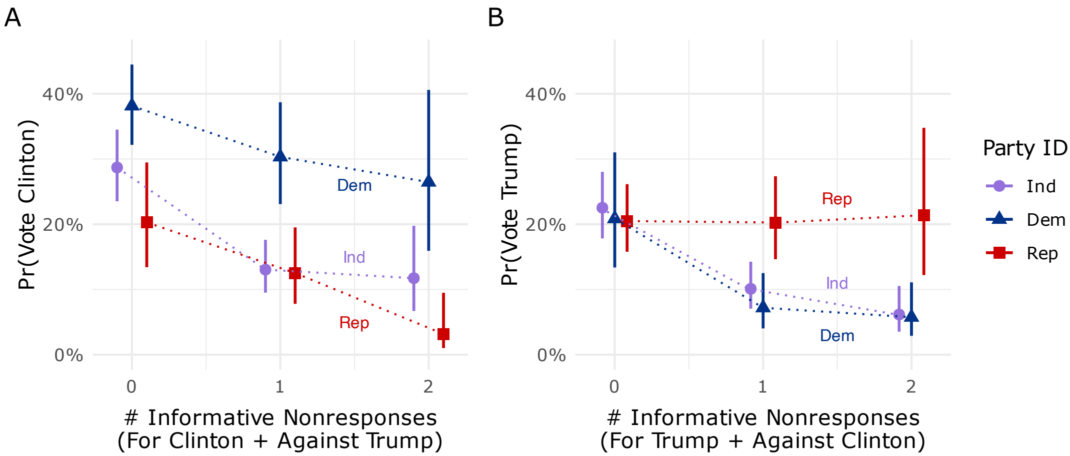

Figure 1. Marginal effects of nonresponse on probability of selecting candidate in 2016 (see Table A1.8).

I test for informative nonresponse using candidate choice and two partisan-congruent prompts. As shown in Equations 2 and 3, nonresponse is defined as the sum of two binary indicators making whether each open-ended response was blank. Among Democrats, for example, when asked what would make them vote for Clinton, only 21 percent have a blank response, while 86 percent do so when asked about voting for Trump. Republicans show nearly symmetrical patterns (see Table A1.1).

The measure is negatively coded, so more blank responses suggest lower support. For simplicity, the measures are labeled with a single candidate but the equations capture both withheld support for the named candidate and unexpressed opposition to their partisan rival—making them directional signals of partisan alignment rather than pure measures of candidate affect. All models in Study 1 are limited to registered voters. Controls include party identification, ideology, racial resentment, hostile sexism, authoritarianism, education, age, gender, race, income, political attention, and survey mode (face-to-face or web). To account for affective polarization, I control for the difference in feeling thermometer ratings, calculated as Trump minus Clinton (hereafter, *candidate affect gap*).

$$\text{Informative Nonresponse}_{\text{Clinton}} = \mathbb{I}[\text{nchar}(\textit{For Clinton}) = 0] + \mathbb{I}[\text{nchar}(\textit{Against Trump}) = 0] \quad (2)$$
$$\text{Informative Nonresponse}_{\text{Trump}} = \mathbb{I}[\text{nchar}(\textit{For Trump}) = 0] + \mathbb{I}[\text{nchar}(\textit{Against Clinton}) = 0] \quad (3)$$

Figure 1 Panel A shows that more nonresponse to the "for Clinton" and "against Trump" questions is associated with about a 14 percentage point drop in support for Clinton among Republicans and Independents, and an 11-point drop among Democrats. Figure 1 Panel B shows that more nonresponse on the "for Trump" and "against Clinton" questions is associated with an 18 percentage point decrease in support for Trump among Democrats and a 19-point decrease among Independents. Notably, Democrats and Independents at zero nonresponse are indistinguishable from Republicans in predicted probability of supporting Trump. Among Republicans, the marginal effect of nonresponse is near zero.

Both informative nonresponse measures show substantial correlations with the candidate affect gap ($|r| \approx 0.67$). Because the models control for this difference in affect, nonresponse can plausibly be interpreted as a form of latent or additional partisan and candidate affect not fully captured by the feeling thermometers and other covariates. Similar results hold when predicting candidate choice in the 2020 and 2024 elections using all nonresponse measures from 2016 (see Section A1.5).

5.0.2 Study 1b: Expressive Alignment on Candidate Choice

In a second set of tests, I investigate whether the number of characters recorded in response to the four ANES candidate evaluation items can predict candidate choice in 2016, 2020 and 2024. As shown in Equations 4–6, I calculate the sum of the transformed character counts for each partisan-congruent pair, take their difference, and divide by two, yielding a scale from –1 to +1.

I refer to this measure as *Expressive Alignment* to reflect four key features. First, the scale is derived from respondents' written or spoken expression, rather than closed-ended survey items. Second, drawing on correlations and principal component analysis (see Section A1.6), I show that the text metadata meaningfully reflect both partisan identification and affect toward the candidates. Like the informative nonresponse measures, Expressive Alignment correlates strongly with individual candidate feeling thermometers ($|r| \approx 0.77$), the difference between Trump and Clinton feeling thermometers ($r = 0.85$), as well as partisan identification ($r = 0.71$). Third, unlike these direct attitudinal measures, Expressive Alignment is inferred from behavior and may capture implicit or less consciously held sentiments. Fourth, the term can also include broader coalitional alignments beyond partisanship—such as divisions that are issue-based, intraparty, or identity-based—making the measure generalizable to contexts where partisan cues are weak or structured differently. In plots, the scale is labeled at either end with "Clinton+" and "Trump+," but the underlying measure reflects a composite of both support for and opposition to the candidates, their parties and associated ideologies.

Although Expressive Alignment correlates highly with the Trump–Clinton feeling thermometer difference, it captures additional behavioral information not reducible to warmth alone. The measure reflects both affect and openness versus closedness, that is whether respondents are willing to even consider and articulate reasons for or against a candidate once prompted. Substantial residual variation (about 28 percent) remains after accounting for the thermometers, and regression results show that Expressive Alignment in 2016 continues to predict candidate choice in later cycles even when the thermometer difference, ideology, and other covariates are held constant. Moreover, interactions with Party ID indicate that the measure helps identify cross-pressured or persuadable voters whose expressive behavior diverges from their partisan labels.

$$\text{Trump Evaluation} = T(\textit{For Trump}) + T(\textit{Against Clinton}) \tag{4}$$

$$\text{Clinton Evaluation} = T(\textit{For Clinton}) + T(\textit{Against Trump}) \tag{5}$$

$$\text{Expressive Alignment} = \frac{\text{Trump Evaluation} - \text{Clinton Evaluation}}{2} \tag{6}$$

Figure 2, Panels A and B, show the marginal effects of Expressive Alignment on 2016 candidate choice, controlling for the same covariates as in Study 1a. Expressive Alignment predicts substantial shifts in the likelihood of supporting a candidate, particularly among *out-party* members. As shown in Panel A, Republicans with high Expressive Alignment toward Clinton are statistically indistinguishable from Democrats in their predicted probability of voting for her—increasing by roughly 40 percentage points as Expressive Alignment moves from +1 (Trump-aligned) to −1 (Clinton-aligned), and by about 30 percentage points among Independents. Likewise, in Panel B, Democrats with strong alignment toward Trump resemble Republicans, with the shift from −1 to +1 corresponding to an increase of about 59 points in the probability of supporting him, and about 43 points among Independents. In contrast, the slopes for *co-partisans* remain relatively flat, largely reflecting overlap with the feeling thermometer measure included in the model. Overall, Expressive Alignment captures unmeasured variation in candidate choice not explained by ideology or feeling thermometer ratings and continues to predict large shifts in candidate choice through 2020 and 2024 (see Sections A1.7 and A1.8).

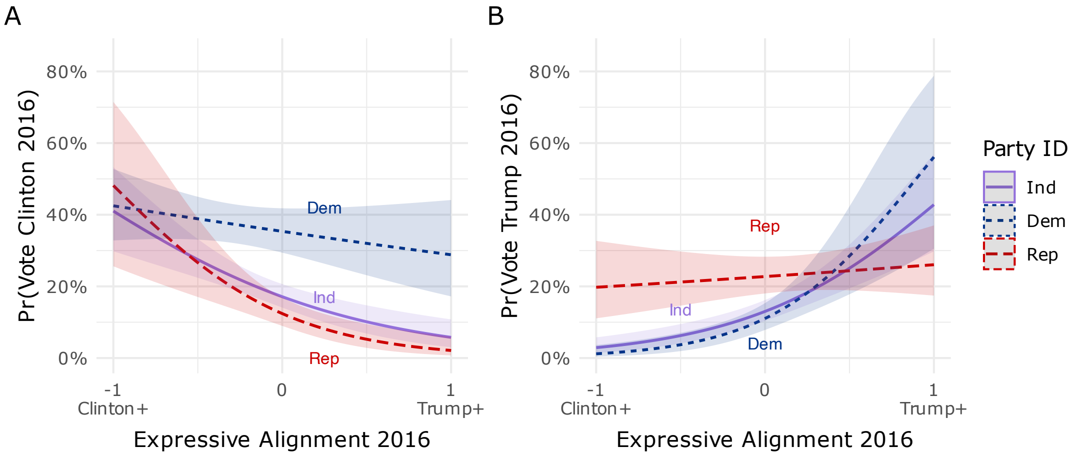

Figure 2. Marginal effects of Expressive Alignment in 2016 on candidate choice in 2016, by party identification (see Table A1.12).

5.0.3 *Study 2: Text as Validated Turnout*

In Study 2, I test whether text as behavior can also improve prediction of validated turnout. Looking at the 2016 presidential election, Kim, Alvarez, and Ramirez (2020) report they "do not find that many demographic variables were closely associated with turnout in 2016" (979). Common predictors of turnout—such as income, education, or political attention—are only rough proxies for the likelihood to vote (Kim, Alvarez, and Ramirez 2020). An education variable, for instance, does not reflect school quality or an interest in learning about the world, and many college graduates spend more time watching entertainment programming than news. In contrast, open-ended prompts may yield more granular insight into the intensity of individual political engagement. Also, because open-ended prompts require more effort, they may reduce response bias relative to closed-ended items where selecting a more extreme response (e.g., overstating political interest) carries little additional cost.

This analysis draws on a separate 2016 ANES battery in which respondents were asked to name and prioritize the "most important problems in America" (henceforth, MIP) across four open-ended prompts. As these questions are not partisan, I sum transformed character counts across all four responses:

$$\text{Expressive Engagement} \quad = \quad T(\textit{Problem 1}) + T(\textit{Problem 2}) + T(\textit{Problem 3}) + T(\textit{Problem 4}) \tag{7}$$

I refer to this measure as *Expressive Engagement* to highlight that, while correlated with civic and political engagement, it is rooted in a willingness to articulate political concerns. Principal component analysis shows that Expressive Engagement loads most heavily on a dimension that distinguishes articulation (e.g., verbosity) from concrete political behaviors (e.g., donating, volunteering). It also loads more weakly on a second dimension representing conventional political engagement, such as political interest and knowledge (see Section A2.2). These results suggest that Expressive Engagement reflects both established forms of participation and a more symbolic mode of political involvement through communication.

Figure 3 shows the predicted probability of validated turnout as a function of Expressive Engagement among registered voters controlling for sex, education, age, race, income, political attention, survey mode, party identification, and self-reported likelihood to vote (for turnout by registration status, see Section A1.2). Because self-reported likelihood to vote is by far the strongest predictor of turnout, its inclusion makes this a particularly stringent test of the Expressive Engagement scale. Panel A shows that, after adjusting for covariates, moving from zero to nearly four on Expressive Engagement is associated with a significant and meaningful 16 percentage point increase in the predicted probability of voting in 2016. Drawing on the panel structure of the data, Panel B shows the same change in Expressive Engagement is associated with a roughly 18 percentage point increase in the likelihood of validated turnout four years later, though estimated with greater uncertainty. Secondary analyses using Expressive Engagement in 2020 on turnout in 2020 and Expressive Alignment in 2016 on turnout in 2016 and 2020 are reported in the Appendix (see Sections A2.4 and A2.5).[3]

3. One possible concern is that Expressive Engagement predicts willingness to provide matchable identifying information rather than turnout itself. The ANES supplies both a validated-turnout indicator and a match probability. To address this concern, I incorporate linkage uncertainty directly into the dependent variable by defining the outcome as their product. Results suggest the relationship is not driven by matchability, and are substantively unchanged when using binary validated turnout, match-probability-weighted models, or restricting the sample to high-certainty matches.

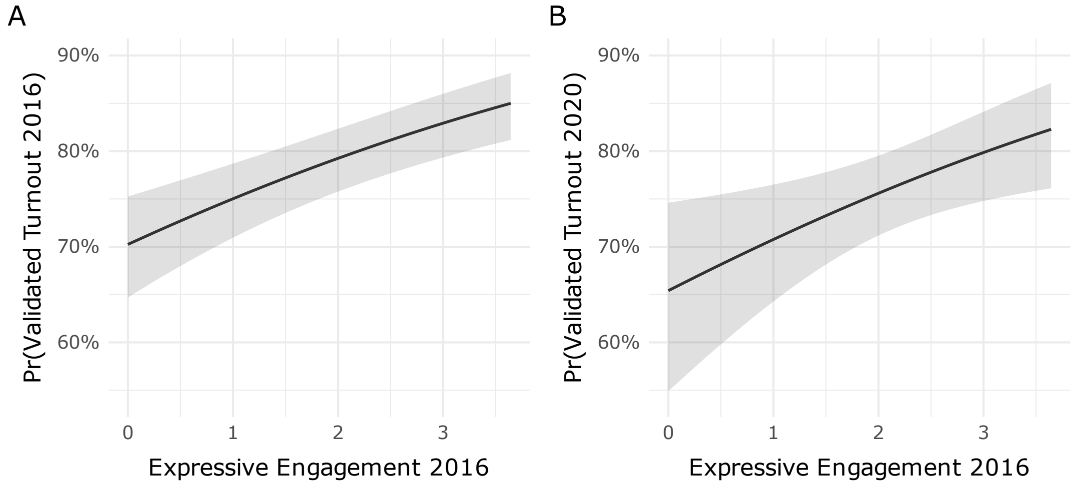

Figure 3. Marginal effects of Expressive Engagement in 2016 on validated turnout incorporating match probability in (A) 2016 and (B) 2020 using quasibinomial models (see Table A2.18).

5.0.4 Study 3: Text as Behavioral Outcome

Measuring real-world behavior in survey research is notoriously difficult. Researchers often rely on games or incentivized tasks to elicit sincere preferences, but these methods can be costly, artificial, or limited in external validity (McDermott 2002; Winking and Mizer 2013; Findley, Kikuta, and Denly 2021). In Study 3, I draw on the 2016 Afrobarometer to test whether text metadata can also serve as a behavioral *outcome* within a traditional survey format (for more on the Afrobarometer, see Section 4.2).

Features of text such as character counts and nonresponse plausibly reflect underlying features of human affect and cognition that operate largely independently of any specific language. I begin with a battery of ten questions measuring support for democratic governance, such as whether respondents prefer to "choose leaders through elections versus other methods." These ten items are combined into a single "Importance of Democracy" scale, which I validate using the pooled character count from three open-ended prompts asking "What, if anything, does 'democracy' mean to you?"

Figure 4 tests whether responses to the ten democracy questions predict the length of responses to the prompt, "What does democracy mean to you?" (controlling for gender, education, age, language, race, and income proxy).[4] I use raw character counts for ease of interpretation, though results are robust to the transformations described in Section 4.3.

For the purpose of validating the instrument across languages, the key concern is whether the *slopes* show that the Importance of Democracy scale is associated with more open-ended expression. As shown in Figure 4, the slopes are all positive, significantly different from zero, and substantively similar. The French-language slope is statistically steeper than the English baseline, but the magnitude of the difference is modest, and all three languages exhibit similarly sized positive associations between valuing democracy and response length. The intercepts also differ modestly but significantly across the three languages, suggesting baseline differences in verbosity.

In sum, these results suggest that individuals who report valuing democracy in closed-ended questions also tend to "pay costs" by offering longer responses to related open-ended prompts. More broadly, Figure 4 suggests character counts can serve as valid behavioral measures across languages, even after transcription and translation.

4. I drop 18 percent of subjects who reported, "Did not understand the word [democracy] or question, even in local language." Results are robust to their inclusion.

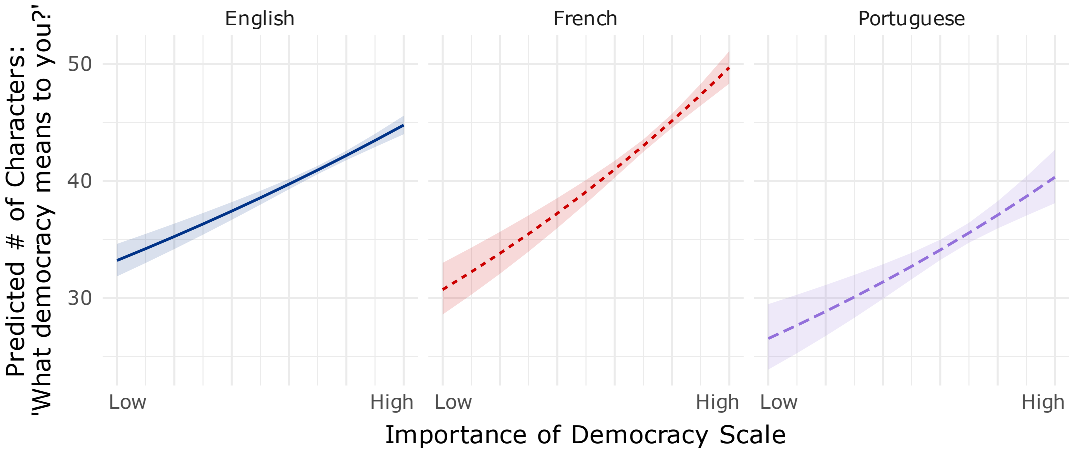

Figure 4. Marginal effects of Importance of Democracy Scale on number of characters written in response to open-ended prompts asking 'What democracy means to you,' by language. The negative binomial model includes controls for gender, education, age, race and income-proxy (see Table A3.26).

6. Discussion

While these studies show metadata from open-ended prompts can serve as useful measures of attitudes and behavior, at least five important questions remain. First, although the results above suggest that certain types of prompts are closely related to specific attitudes and behaviors, more work is needed to map which kinds are most effective for measuring particular beliefs, dispositions, and actions. Second, measurement challenges remain. These include how to address skewed distributions (e.g., IHS, log, truncation) and how best to model zero-character responses. Some forms of missing data, such as attrition or saying "Don't know," may also constitute types of informative nonresponse. Although I apply this logic to the Afrobarometer data (see Section A3.4), a fuller treatment falls outside the scope of this paper.

Third, more research is needed on detecting fraud in open-ended survey items. As AI tools proliferate, researchers face new forms of automated or AI-assisted deception (Westwood 2025). For example, Veselovsky, Ribeiro, and West (2023) capture keystrokes, including copy–paste actions, to identify synthetic text production. Other behavioral traces, such as typing speed, keystroke regularity, and mouse movements can flag automation, though such signals can themselves be forged (Westwood 2025). These *process-based* indicators differ from the *content-level* mischief described by Lopez and Hillygus (2018), where respondents intentionally provide absurd answers for amusement.

Text as behavior methods suggest a complementary approach: looking for *discrepant-effort* patterns such as respondents who report high political interest yet contribute minimal text to related prompts like the "most important problem" questions. Such mismatches could flag inattentive or mischievous respondents who are difficult to detect through content alone. More broadly, incentives can enhance response quality (Li 2023), and some subtle behavioral measures, like nonresponse, may prove more resistant to forms of AI manipulation.

Fourth, I do not address how to integrate text as behavior with other types of computational text analysis, such as machine learning methods. Finally, although the analyses here suggest that text metadata can be applied across languages, cultures, and survey modes, more research is needed to assess their generalizability. Cultural norms about speech, silence, or verbosity may affect the interpretability of character counts. The transformations used here, such as deskewing and normalization by mode, usefully merged data across modes but the cognitive dynamics of writing versus speaking may, nevertheless, introduce meaningful variation in metadata signals (Benoit, Munger, and Spirling 2019). Topics beyond politics should also be considered. Further research should explore how cultural context, survey mode, and text source influence the validity of metadata as behavioral measures.

7. Conclusion

Quantifying human attitudes and behaviors remains a fundamental challenge in social science. This study proposes extending text as data methods by treating text as behavior. Specifically, expression is sufficiently cognitively, affectively and temporally demanding that it should be understood not only as a means of communication, but also as a form of action. As shown in prior work and in the analyses presented here, metadata from open-ended responses can offer insight into underlying mental and emotional states, even when those states are not fully accessible to respondents themselves. Measuring human behavior will always be difficult, but the findings here suggest that treating expression as effortful provides social scientists an additional tool to reveal preferences, commitments, and underlying dispositions. As Gloria Steinem once said, "I don't like writing. I like having written" (1976).

Supplementary Material

The Online Appendix is available as supplementary material.

Data Availability Statement

Replication code for this article has been published in the Political Analysis Harvard Dataverse at https://doi.org/10.7910/DVN/HKBZNG (Wasow 2026).

Author ORCiD

Omar Wasow: https://orcid.org/0000-0002-1104-4610

Acknowledgements

I thank Risa Gelles-Watnick and Christina Im for excellent research assistance. I thank Andrew Little, Elena Llaudet, Daniel Masterson, John Konicki, Zach Hertz, Alexander Agadjanian, and three anonymous reviewers for extremely helpful feedback. AI tools used to proofread manuscript and prepare replication archive.

Conflict of Interest

The author declares no conflicts of interest.

Ethical Standards

This study uses publicly available secondary data and did not require IRB approval.

References

Alsaedi, Nasser, Pete Burnap, and Omer Rana. 2017. "Can We Predict a Riot? Disruptive Event Detection Using Twitter." *ACM Trans. Internet Technol* 17 (2).

American National Election Studies. 2017. "ANES 2016 Time Series Study Full Release." Dataset and documentation. Available at: https://www.electionstudies.org.

———. 2021. "ANES 2020 Time Series Study Full Release." Dataset and documentation, February 10, 2022 version.

———. 2025. "ANES 2024 Time Series Study Preliminary Release." Dataset and documentation, April 30, 2025 version.

Aronson, Elliot, and J Merrill Carlsmith. 1968. "Experimentation in Social Psychology." *The Handbook of Social Psychology* 2 (2): 1–79.

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.

Berinsky, Adam J. 2008. "Survey Non-Response." *The SAGE Handbook of Public Opinion Research*, 309–21.

———. 2013. *Silent Voices: Public Opinion and Political Participation in America*. Princeton University Press.

Bernhard, Rachel, and Sean Freeder. 2020. "The More You Know: Voter Heuristics and the Information Search." *Political Behavior* 42 (2): 603–23.

Bisgaard, Martin. 2019. "How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning." *American Journal of Political Science* 63 (4): 824–39.

Butler, Daniel M, and David E Broockman. 2011. "Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators." *American Journal of Political Science* 55 (3): 463–77.

Campbell, Donald T, and Julian C Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Houghton Mifflin Company.
Cavaillé, Charlotte, Daniel L. Chen, and Karine Van der Straeten. 2024. "Who Cares? Measuring Differences in Preference Intensity." *Political Science Research and Methods*, 1–17.
Crichton, Michael. 1984. "The Mousetrap: A Tale of Computer Crime." *Life Magazine* 7 (1): 116–26.
Danescu-Niculescu-Mizil, Cristian, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. "A Computational Approach to Politeness with Application to Social Factors."
Day, Martin V, Susan T Fiske, Emily L Downing, and Thomas E Trail. 2014. "Shifting Liberal and Conservative Attitudes Using Moral Foundations Theory." *Personality and Social Psychology Bulletin* 40 (12): 1559–73.
Denscombe, Martyn. 2008. "The Length of Responses to Open-Ended Questions: A Comparison of Online and Paper Questionnaires in Terms of a Mode Effect." *Social Science Computer Review* 26 (3): 359–68.
Eady, Gregory, Frederik Hjorth, and Peter Thisted Dinesen. 2022. "Do Violent Protests Affect Expressions of Party Identity? Evidence from the Capitol Insurrection." *American Political Science Review*, 1–7.
Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. 2017. "'User's Guide and Codebook for the ANES 2016 Time Series Voter Validation Supplemental Data.'" Technical report, Princeton University.
Findley, Michael G., Kyosuke Kikuta, and Michael Denly. 2021. "External Validity." *Annual Review of Political Science* 24 (1): 365–93.
Gillion, Daniel Q. 2016. *Governing with Words: The Political Dialogue on Race, Public Policy, and Inequality in America*. Cambridge University Press.
Graham, Steve. 2018. "A Revised Writer(s)-Within-Community Model of Writing." *Educational Psychologist* 53 (4): 258–79.
Grimmer, Justin, Margaret E Roberts, and Brandon M Stewart. 2022. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press.
Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Religion and Politics*. New York: Pantheon Books.
Hayes, John R. 1996. "A New Framework for Understanding Cognition and Affect in Writing." In *The Science of Writing: Theories, Methods, Individual Differences and Applications*, edited by C. Michael Levy and Sarah Ransdell, 6–44. Lawrence Erlbaum Associates.
Horwitz, Rachel, Frauke Kreuter, and Frederick Conrad. 2017. "Using Mouse Movements to Predict Web Survey Response Difficulty." *Social Science Computer Review* 35 (3): 388–405.
Jacobi, Tonja, and Dylan Schweers. 2017. "Justice, Interrupted: The Effect of Gender, Ideology, and Seniority at Supreme Court Oral Arguments." *Va. L. Review* 103: 1379.
Jann, Ole, and Christoph Schottmüller. 2024. "Political Debate on Social Media: Theory and Evidence." https://schottmueller.github.io/papers/twitter/twitter.pdf.
Jenke, Libby, Kirk Bansak, Jens Hainmueller, and Dominik Hangartner. 2021. "Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments." *Political Analysis* 29 (1): 75–101.
Karpowitz, Christopher F, Tali Mendelberg, and Lee Shaker. 2012. "Gender Inequality in Deliberative Participation." *American Political Science Review* 106 (3): 533–47.
Kellogg, Ronald T. 1999. *The Psychology of Writing*. Oxford University Press.
Kim, Seo-young Silvia, R. Michael Alvarez, and Christina M. Ramirez. 2020. "Who Voted in 2016? Using Fuzzy Forests to Understand Voter Turnout." *Social Science Quarterly* 101 (2): 978–88.
Kraft, Patrick W. 2023. "Women Also Know Stuff: Challenging the Gender Gap in Political Sophistication." *American Political Science Review*, 1–19.
Kramer, Adam DI, Jamie E Guillory, and Jeffrey T Hancock. 2014. "Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks." *Proceedings of the National*

Academy of Sciences 111 (24): 8788–90.
Kuo, Alexander, Neil Malhotra, and Cecilia Hyunjung Mo. 2017. "Social Exclusion and Political Identity: The Case of Asian American Partisanship." *The Journal of Politics* 79 (1): 17–32.
Laaksonen, Seppo, and Ray Chambers. 2006. "Survey Estimation Under Informative Nonresponse with Follow-up." *Journal of Official Statistics* 22 (1): 81–95.
Li, Yimeng. 2023. "Incentivizing Open-Ended Survey Responses." Presented at Polmeth 2023, July 6, 2023 at Stanford University.
Longford, Nicholas T. 2007. "Missing Data." In *Handbook of Multilevel Analysis*, edited by Jan de Leeuw and Erik Meijer, 379–401. Berlin: Springer.
Lopez, Jesse, and D. Sunshine Hillygus. 2018. "Why so Serious?: Survey Trolls and Misinformation." https://doi.org/10.2139/ssrn.3131087.
Lowande, Kenneth, and Andrew Proctor. 2020. "Bureaucratic Responsiveness to LGBT Americans." *American Journal of Political Science* 64 (3): 664–81.
Maloney, Elizabeth. 2021. "The Gender Wage Gap: A Product of Misogyny and Gender Norms." Online at https://www.economics.uci.edu/files/docs/2021/gradjobmarket/maloney_elizabeth.pdf.
Matthews, Peter Hugoe. 2003. *Linguistics: A Very Short Introduction*. Oxford University Press.
McDermott, Rose. 2002. "Experimental Methodology in Political Science." *Political Analysis* 10 (4): 325–42.
Mosleh, Mohsen, Cameron Martel, Dean Eckles, and David Rand. 2021. "Perverse Downstream Consequences of Debunking: Being Corrected by Another User for Posting False Political News Increases Subsequent Sharing of Low Quality, Partisan, and Toxic Content in a Twitter Field Experiment." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
Mossholder, Kevin W, Randall P Settoon, Stanley G Harris, and Achilles A Armenakis. 1995. "Measuring Emotion in Open-Ended Survey Responses: An Application of Textual Data Analysis." *Journal of Management* 21 (2): 335–55.
Mummolo, Jonathan, and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113 (2): 517–29.
Munger, Kevin. 2017. "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment." *Political Behavior* 39 (3): 629–49.
Panteli, Niki. 2002. "Richness, Power Cues and Email Text." *Information & Management* 40 (2): 75–86.
Patterson, Molly, and Kristen Renwick Monroe. 1998. "Narrative in Political Science." *Annual Review of Political Science* 1: 315–31.
Pennebaker, James W. 1997. "Writing about Emotional Experiences as a Therapeutic Process." *Psychological Science* 8 (3): 162–66.
Pennebaker, James W, Martha E Francis, and Roger J Booth. 2001. "Linguistic Inquiry and Word Count: LIWC 2001." *Mahway: Lawrence Erlbaum Associates* 71.
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): 1064–82.
Rude, Stephanie, Eva-Maria Gortner, and James Pennebaker. 2004. "Language Use of Depressed and Depression-Vulnerable College Students." *Cognition & Emotion* 18 (8): 1121–33.
Sen, Maya, and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19: 499–522.
Steinem, Gloria. 1976. "As Quoted in Pocatellan Captures Flavor of Stanley Basin's History by Joan LaLiberte." *Idaho State Journal*.
Stephens-Davidowitz, Seth. 2014. "The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data." *Journal of Public Economics* 118: 26–40.

Teh, Pin Shen, Andrew Beng Jin Teoh, and Shigang Yue. 2013. "A Survey of Keystroke Dynamics Biometrics." *The Scientific World Journal* 2013 (1): 1–24.

ten Kleij, Frederieke, and Pieter A. D Musters. 2003. "Text Analysis of Open-Ended Survey Responses: A Complementary Method to Preference Mapping." *Food Quality and Preference* 14 (1): 43–52.

Thompson, Clive. 2010. "Clive Thompson on the Death of the Phone Call." *Wired*. https://www.wired.com/2010/07/st-thompson-deadphone/.

Tumasjan, Andranik, Timm O Sprenger, Philipp G Sandner, and Isabell M Welpe. 2010. "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment." In *Fourth International AAAI Conference on Weblogs and Social Media*. Citeseer.

van Zanden, Jan Luiten, Joerg Baten, Marco Mira d'Ercole, Auke Rijpma, and Marcel P. Timmer, eds. 2014. *How Was Life?: Global Well-Being Since 1820*. Paris: OECD Publishing.

Veselovsky, Veniamin, Manoel Horta Ribeiro, and Robert West. 2023. "Artificial Artificial Artificial Intelligence: Crowd Workers Widely Use Large Language Models for Text Production Tasks."

Voigt, Rob, Nicholas P Camp, Vinodkumar Prabhakaran, William L Hamilton, Rebecca C Hetey, Camilla M Griffiths, David Jurgens, Dan Jurafsky, and Jennifer L Eberhardt. 2017. "Language from Police Body Camera Footage Shows Racial Disparities in Officer Respect." *Proceedings of the National Academy of Sciences* 114 (25): 6521–26.

Wasow, Omar. 2026. "Replication Data for: Text as Behavior." Harvard Dataverse, V1. https://doi.org/10.7910/DVN/HKBZNG.

Wasow, Omar, Alex Baron, Marlon Gerra, Katherine Lauderdale, and Han Zhang. 2010. "Can Tweets Kill a Movie? An Empirical Evaluation of the Bruno Effect"." Presented at the workshop on Microblogging at the Conference on Human Computer Interaction (CHI), April 11 2010. Online at https://ssrn.com/abstract=5628071.

Westwood, Sean J. 2025. "The Potential Existential Threat of Large Language Models to Online Survey Research." *Proceedings of the National Academy of Sciences* 122 (47): e2518075122.

Wilson, Timothy D. 2004. *Strangers to Ourselves*. Cambridge, MA: Harvard University Press.

Winking, Jeffrey, and Nicholas Mizer. 2013. "Natural-Field Dictator Game Shows No Altruistic Giving." *Evolution and Human Behavior* 34 (4): 288–93.

Wright, Robert. 2017. *Why Buddhism Is True: The Science and Philosophy of Meditation and Enlightenment*. Simon; Schuster.

Yan, Alan N., and Rachel Bernhard. 2023. "The Silenced Text: Field Experiments on Gendered Experiences of Political Participation." *American Political Science Review*.

Ziegler, Jeffrey. 2022. "A Text-as-Data Approach for Using Open-Ended Responses as Manipulation Checks." *Political Analysis* 30 (2): 289–97.

Appendix

A1. Study 1: Candidate Choice and Party Identification vs Nonresponse and Expressive Alignment

A1.1 Study 1a: Candidate Choice vs Nonresponse with Candidate Evaluation Prompts

Party ID	Candidate Evaluation	# Nonresponse	n	Percent
Democrat	For Clinton	0	1,139	79%
		1	311	21%
	Against Clinton	0	473	33%
		1	977	67%
	For Trump	0	210	14%
		1	1,240	86%
	Against Trump	0	1,262	87%
		1	188	13%
Independent	For Clinton	0	547	40%
		1	820	60%
	Against Clinton	0	909	66%
		1	458	34%
	For Trump	0	587	43%
		1	780	57%
	Against Trump	0	961	70%
		1	406	30%
Republican	For Clinton	0	181	15%
		1	1,050	85%
	Against Clinton	0	1,081	88%
		1	150	12%
	For Trump	0	970	79%
		1	261	21%
	Against Trump	0	550	45%
		1	681	55%

Table A1.1: Nonresponse by Party ID and Individual Candidate Evaluation Prompts

Party ID	Candidate Evaluation	# Nonresponse	n	Percent
Democrat	For Clinton + Against Trump	0	1,070	74%
		1	261	18%
		2	119	8%
	For Trump + Against Clinton	0	138	10%
		1	407	28%
		2	905	62%
Independent	For Clinton + Against Trump	0	515	38%
		1	478	35%
		2	374	27%
	For Trump + Against Clinton	0	539	39%
		1	418	31%
		2	410	30%
Republican	For Clinton + Against Trump	0	151	12%
		1	429	35%
		2	651	53%
	For Trump + Against Clinton	0	903	73%
		1	245	20%
		2	83	7%

Table A1.2: Nonresponse by Party ID and Congruent Candidate Evaluation Prompts

A1.2 Study 1a: Voting Behavior by Registration Status in 2016 and 2020

	Did Not Vote	Voted
Not Registered	524 (85.6%)	88 (14.4%)
Registered	1,036 (28.3%)	2,622 (71.7%)

Table A1.3: Turnout by Registration Status in 2016

	Did Not Vote	Voted
Not Registered	405 (63.9%)	229 (36.1%)
Registered	1,200 (16.2%)	6,185 (83.8%)

Table A1.4: Turnout by Registration Status in 2020

A1.3 Study 1a: Number of character summary statistics for candidate for/against responses by mode and year

	Max		Mean		Median	
Variable	FTF	Web	FTF	Web	FTF	Web
For Clinton	1024	60	59.1	12.6	12	0
Against Clinton	1024	60	76.4	16.2	32	10
For Trump	1024	60	60.6	12.0	0	0
Against Trump	1024	60	89.5	20.6	48	14

Note: Face-to-face: n = 1,180, Web: n = 3,090.

Table A1.5: Number of character summary statistics for candidate for/against responses by survey mode16 (2016)

	Max			Mean			Median		
Variable	Tele	Video	Web	Tele	Video	Web	Tele	Video	Web
For Biden	810	1038	1200	88.2	87.6	46.8	36	47	0
Against Biden	911	851	1199	73.2	64.8	48.6	19	0	0
For Trump	969	462	1198	62.8	54.9	48.1	0	0	0
Against Trump	1200	1277	1217	125.4	135.9	77.9	51	82	29

Note: Tele: n = 139, Video: n = 359, Web: n = 7,782.

Table A1.6: Number of character summary statistics for candidate for/against responses by survey mode16 (2020)

Variable	Max			Mean			Median		
	FTF	Tele	Web	FTF	Tele	Web	FTF	Tele	Web
For Harris	1299	811	4712	84.5	102.0	46.8	20.5	39.5	0
Against Harris	4578	836	2638	92.4	70.4	50.3	0.0	0.0	0
For Trump	1357	440	5167	86.0	58.9	45.8	0.0	0.0	0
Against Trump	2576	756	2379	128.5	127.8	69.1	70.0	63.0	28

Note: Pre-election, Face-to-face: n = 966, Tele: n =76, Web: n = 4,234

Table A1.7: Number of character summary statistics for candidate for/against responses by survey mode16 (2024)

A1.4 Study 1a: Candidate Choice in 2016 vs Nonresponse in 2016

	Dependent variable:			
	Clinton 2016	Trump 2016	Clinton 2016	Trump 2016
	(1)	(2)	(3)	(4)
# Nonresp (For Clinton + Against Trump)	−1.666*		−0.988*	
	(0.198)		(0.220)	
# Nonresp (For Clinton + Against Trump)	−2.489*		−1.108*	
	(0.323)		(0.345)	
# Nonresp (For Trump + Against Clinton)		−1.697*		−0.951*
		(0.202)		(0.230)
# Nonresp (For Trump + Against Clinton)		−2.671*		−1.488*
		(0.295)		(0.325)
Party: Democrats	0.968*	−0.870*	0.426*	−0.098
	(0.146)	(0.245)	(0.163)	(0.297)
Party: Republican	−0.538*	0.281	−0.459	−0.121
	(0.245)	(0.152)	(0.279)	(0.171)
Ideology (1-7)	−0.243*	0.411*	−0.107*	0.263*
	(0.044)	(0.049)	(0.050)	(0.054)
Feeling Therm: Trump - Clinton			−0.029*	0.034*
			(0.002)	(0.002)
Racial Resentment	−0.433*	0.443*	−0.198*	0.036
	(0.060)	(0.064)	(0.068)	(0.076)
Hostile Sexism	−0.170*	0.160*	−0.030	0.029
	(0.066)	(0.068)	(0.074)	(0.078)
Authoritarianism	−0.082	0.053	−0.080	−0.051
	(0.048)	(0.049)	(0.053)	(0.056)
Education	0.028	−0.014	0.072*	0.051
	(0.027)	(0.028)	(0.030)	(0.032)
Age (yrs)	0.098*	0.085*	0.078*	0.111*
	(0.017)	(0.017)	(0.018)	(0.019)
Female	0.016	0.250*	−0.169	0.254*
	(0.111)	(0.112)	(0.123)	(0.125)
Income	0.031*	0.012	0.036*	0.025*
	(0.007)	(0.008)	(0.008)	(0.009)
Political Attention	0.723*	0.244	0.800*	−0.106
	(0.209)	(0.215)	(0.230)	(0.243)
Mode: Web	0.050	0.172	−0.241	−0.061
	(0.118)	(0.118)	(0.132)	(0.134)
Party: Democrat	0.798*		0.640*	
	(0.267)		(0.297)	
Party: Democrat	0.715		0.569	
	(0.431)		(0.475)	
Party: Republican	−0.146		0.413	
	(0.379)		(0.420)	
Party: Republican	−1.388*		−0.942	
	(0.695)		(0.722)	
Party: Democrat		−0.159		−0.273
		(0.412)		(0.469)
Party: Democrat		−0.247		0.025
		(0.499)		(0.552)
Party: Republican		0.945*		0.937*
		(0.273)		(0.310)
Party: Republican		1.372*		1.544*
		(0.416)		(0.475)
Constant	0.861	−5.737*	−1.987*	−4.832*
	(0.547)	(0.652)	(0.615)	(0.730)
Feeling Thermometer?	No	No	Yes	Yes
Observations	3,320	3,320	3,289	3,289
Log Likelihood	-1,168.468	-1,128.254	-982.553	-919.865
Akaike Inf. Crit.	2,390.935	2,310.509	2,021.106	1,895.729

Note: Race variable included in models but omitted above for space.
*$p < 0.05$

Table A1.8: Candidate Choice vs Nonresponse (Discrete)

A1.5 Study 1a: Candidate Choice in 2020 and 2024 vs All Nonresponse in 2016

In Equation 8, I subtract the results of Equation 3 (Informative Nonresponse$_{\text{Trump}}$) from the results of Equation 2 (Informative Nonresponse$_{\text{Clinton}}$) to create a single Nonresponse Scale. Because each Informative Nonresponse measure is negatively coded—meaning higher values indicate more negative views of the candidate—subtracting Trump from Clinton yields a scale where negative values indicate more pro-Clinton (or liberal) sentiment and positive values indicate more pro-Trump (or conservative) sentiment. The scale therefore runs from –2 (most pro-Clinton) to +2 (most pro-Trump). As the possible number of nonresponses for each pair of prompts runs from 0 to 2, for ease of interpretation I do not normalize the scale to –1 to +1 as I do with Expressive Alignment.

$$\text{Nonresponse Scale 2016} \quad = \quad \text{Informative Nonresponse}_{\text{Clinton}} - \text{Informative Nonresponse}_{\text{Trump}} \tag{8}$$

In Figure A1.1, Panels A and B, we see that moving from –2 to +2 on the Nonresponse Scale in 2016 is associated with substantial changes in the predicted probability of selecting a candidate among *outparty* members, but relatively little change among *copartisans*.

For example, in Panel A, among Republicans, the Nonresponse Scale predicts an approximately 67 percentage point decrease in the probability of prefering Biden in 2020 while, among Democrats, the predicted decrease is only about 7 percentage points. In Panel B, among Democrats, moving from –2 to +2 on the Nonresponse Scale predicts a roughly 55 percentage point increase in the probability of selecting Trump, but only a 16 percentage point increase among Republicans. Importantly, because these models control for ideology, party ID, and the difference in feeling thermometer ratings between Trump and Clinton, the expressive behavior measured by the Nonresponse Scale appears to capture meaningful residual variation that standard attitudinal measures do not fully explain.

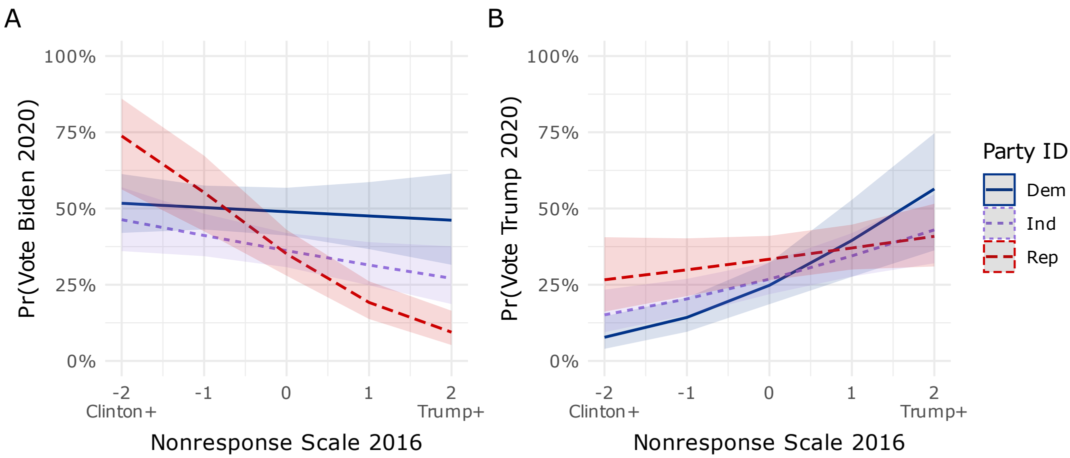

Figure A1.1. Marginal effects of Nonresponse Scale in 2016 on probability of selecting candidate in 2020, by party identification (see Table A1.9).

In Figure A1.2, a similar pattern appears for predicting 2024 candidate choice using the 2016 Nonresponse Scale. In both Panels A and B, increases on the Nonresponse Scale are associated with large shifts in candidate preference, especially among Republicans and Independents. Panel A shows that the probability of a 2024 Harris vote drops sharply from 0.88 to 0.24 among Republicans as nonresponse increases from –2 to +2, a 64 percentage point swing. Among Independents, the drop is nearly 40 percentage points (from 0.76 to 0.37). Among Democrats, the shift is more modest, falling by 20 points (0.73 to 0.53). Panel B shows a corresponding pattern in Trump support: the probability of preferring Trump in 2024 increases from 0.12 to 0.60 among Republicans, and from

	Dependent variable:	
	Vote Biden 2020	Vote Trump 2020
	logistic	*logistic*
	(1)	(2)
Nonresponse Scale	−0.211*	0.361*
	(0.098)	(0.109)
Party: Democrats	0.526*	−0.105
	(0.171)	(0.206)
Party: Republican	−0.045	0.312
	(0.183)	(0.176)
Ideology (1-7)	−0.095	0.209*
	(0.050)	(0.055)
Feeling Therm: Trump - Clinton	−0.019*	0.020*
	(0.002)	(0.002)
Racial Resentment	−0.284*	0.312*
	(0.067)	(0.075)
Hostile Sexism	−0.151*	−0.061
	(0.073)	(0.078)
Authoritarianism	−0.110*	0.055
	(0.051)	(0.056)
Education	0.041	0.144*
	(0.029)	(0.032)
Age (yrs)	0.073*	0.039*
	(0.018)	(0.019)
Female	0.060	0.244
	(0.120)	(0.126)
Race: Black	0.207	−0.458
	(0.340)	(0.527)
Race: Hispanic	0.073	−0.387
	(0.333)	(0.413)
Race: Native American	0.547	0.499
	(0.823)	(0.995)
Race: Other	0.089	−0.092
	(0.407)	(0.471)
Race: White	0.319	0.121
	(0.294)	(0.335)
Income	0.035*	0.010
	(0.008)	(0.009)
Political Attention	0.588*	−0.036
	(0.229)	(0.244)
Mode: Web	0.041	0.037
	(0.132)	(0.138)
Nonresponse Scale × Dem	0.155	0.323
	(0.127)	(0.187)
Nonresponse Scale × Rep	−0.614*	−0.200
	(0.174)	(0.136)
Constant	−0.958	−5.405*
	(0.602)	(0.699)
Observations	2,573	2,573
Log Likelihood	-999.671	-897.595
Akaike Inf. Crit.	2,047.342	1,843.190
Note:		*$p < 0.05$

Table A1.9: Candidate Choice vs Nonresponse Scale

0.12 to 0.52 among Independents and from from 0.17 to 0.42 among Democrats. Notably, in Panel B the slopes do not differ significantly, suggesting a weakening of the boundary between expressive outparty and copartisan behavior (also see the interaction terms in Table A1.10 Column 2).

As with all Study 1 analyses, these models control for ideology, affective polarization, racial resentment, and other key attitudinal predictors. This suggests that the expressive behavior captured by the Nonresponse Scale reflects a durable dimension of political behavior that continues to shape candidate choice beyond what standard measures can explain.

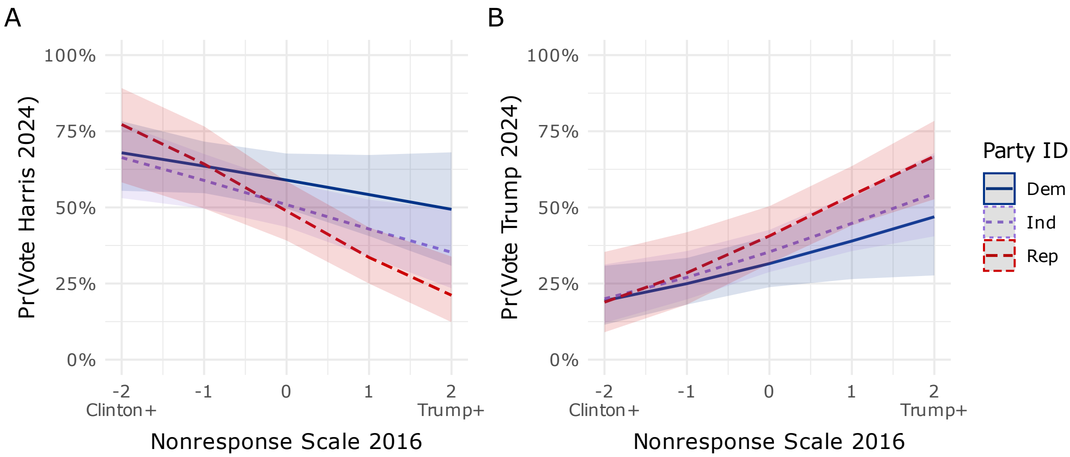

Figure A1.2. Marginal effects of Nonresponse Scale in 2016 on probability of selecting candidate in 2024, by party identification (see Table A1.10).

	Dependent variable:	
	Vote Harris 2024	Vote Trump 2024
	logistic	*logistic*
	(1)	(2)
Nonresponse Scale	−0.322*	0.392*
	(0.123)	(0.127)
Party: Democrats	0.326	−0.173
	(0.205)	(0.211)
Party: Republican	−0.084	0.222
	(0.209)	(0.207)
Ideology (1-7)	−0.180*	0.239*
	(0.065)	(0.066)
Feeling Therm: Trump - Clinton	−0.023*	0.024*
	(0.002)	(0.002)
Racial Resentment	−0.307*	0.265*
	(0.086)	(0.089)
Hostile Sexism	−0.225*	0.060
	(0.094)	(0.096)
Authoritarianism	−0.129*	0.119
	(0.063)	(0.065)
Education	0.080*	−0.045
	(0.037)	(0.038)
Age (yrs)	0.108*	−0.107*
	(0.024)	(0.025)
Female	−0.040	−0.003
	(0.151)	(0.152)
Race: Black	0.790	−0.899
	(0.444)	(0.488)
Race: Hispanic	0.351	−0.021
	(0.414)	(0.431)
Race: Native American	−0.090	−0.023
	(1.102)	(1.095)
Race: Other	0.564	−0.525
	(0.519)	(0.532)
Race: White	0.711*	−0.646
	(0.358)	(0.368)
Income	0.019	−0.015
	(0.010)	(0.010)
Political Attention	0.509	−0.396
	(0.297)	(0.301)
Mode: Web	−0.079	0.174
	(0.168)	(0.169)
Nonresponse Scale x Dem	0.128	−0.067
	(0.167)	(0.184)
Nonresponse Scale x Rep	−0.312	0.147
	(0.192)	(0.184)
Constant	−0.251	−0.664
	(0.768)	(0.792)
Observations	1,983	1,983
Log Likelihood	-654.032	-634.385
Akaike Inf. Crit.	1,356.065	1,316.770
Note:		*$p < 0.05$

Table A1.10: Candidate Choice vs Nonresponse Scale

A1.6 Study 1b: Interpreting Expressive Alignment

The 2016 ANES asked four open-ended questions about the two main presidential candidates, eliciting reasons to vote for or against supporting Hillary Clinton and Donald Trump. As discussed in Equation 4, I combine the number of characters written in response to each question into a scale I call *Expressive Alignment*. Conceptually, this measure builds on the idea that expressive verbosity in political contexts reflects both affective intensity and motivation to engage with partisan meaning-making. Longer open-ended responses suggest a willingness to elaborate emotional and evaluative stances toward political figures.

In contrast to Feeling Thermometers, Expressive Alignment captures not only warmth versus coolness toward candidates but also openness versus closedness. That is, Expressive Alignment reflects whether individuals are willing to entertain reasons for (or against) a candidate or reject the possibility outright. Because elaboration follows a conditional "Yes" response, the measure conveys both candidate affect and a behavioral measure of cognitive openness: a willingness to engage, explore, or articulate reasons rather than to dismiss them out of hand. Where the feeling thermometers emphasize valence, Expressive Alignment adds a dimension of curiosity or ambivalence on one side, and moral disgust, certainty, or defensive rigidity on the other. In short, Expressive Alignment reflects both whether individuals would even contemplate articulating reasons for or against a candidate, and the intensity of that stance once they do.

Similarly, nonresponses convey an *unwillingness* to even consider a for or against prompt for a favored or disfavored candidate. To further explore what this scale captures, I examine its empirical relationships to other relevant traits. Table A1.11 shows that Expressive Alignment is strongly correlated with affective and ideological variables: it correlates positively with the difference between the feeling thermometers for Trump and Clinton (0.85), the Republican and Democratic parties (0.74), and conservatives and liberals (0.65). It also correlates negatively with the feeling thermometers toward Clinton ($r = -0.78$) and liberals ($r = -0.60$), and positively with feelings toward Trump ($r = 0.77$), party identification ($r = 0.71$), and the Feeling Thermometer gap between conservatives and liberals ($r = 0.65$).

Term	Correlation
Feeling Therm: Trump - Clinton	0.847
Feeling Therm: Clinton	-0.783
Feeling Therm: Trump 2016	0.774
Feeling Therm: Reps - Dems	0.744
Party ID (Strong Dem = 1, Strong Rep = 7)	0.710
ft_con_lib	0.649
Ideology (1-7)	0.567
Racial Resentment	0.558
Hostile Sexism	0.293
Authoritarianism	0.232

Table A1.11: Select Correlates of Expressive Alignment

Figure A1.3 visualizes these relationships using a Principal Component Analysis (PCA). The PCA shows that Expressive Alignment loads heavily on the first principal component, which captures a liberal/Clinton to conservative/Trump axis defined by party identification, ideology, and candidate feeling thermometers. A secondary component, defined primarily by authoritarianism and sexism, suggests a moral traditionalism dimension largely orthogonal to candidate affect and partisanship. These results indicate that greater Expressive Alignment is strongly associated with ideological and partisan alignment, while remaining distinct from the moral and hierarchical values captured in Dimension 2. Taken together, these findings validate Expressive Alignment as a text-based indicator of partisan motivation and ideological intensity.

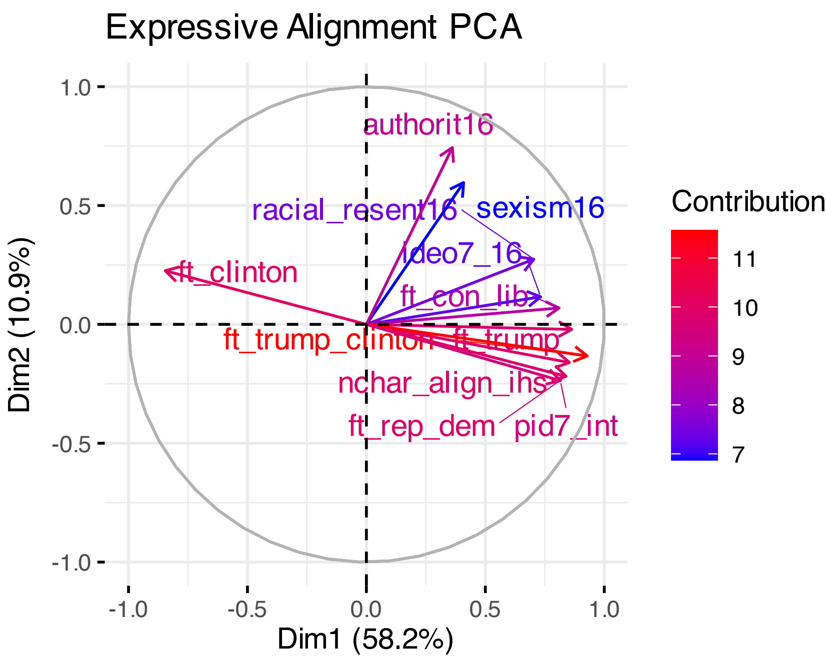

Figure A1.3. Principal Component Analysis shows that Expressive Alignment (nchar_align_ihs) loads heavily on the first principal component, which captures a liberal/Clinton to conservative/Trump axis. A secondary component, defined primarily by authoritarianism and sexism, suggests a moral traditionalism dimension orthogonal to partisanship. These results indicate that higher Expressive Alignment is strongly associated with ideological and partisan alignment, but distinct from the moral/hierarchical values captured in Dimension 2.

It is worth noting that Expressive Alignment correlates so strongly with the candidate feeling thermometer differential (r = 0.85) that including both variables in the same model creates a hard test for Expressive Alignment. The persistence of significant associations for Expressive Alignment suggests that, despite this overlap, it is capturing a distinct aspect of candidate evaluation. While feeling thermometers measure explicit affect or declared preferences toward candidates, Expressive Alignment appears to reflect a more implicit or latent form of affect.

	Dependent variable:	
	Vote Clinton 2016	Vote Trump 2016
	logistic	*logistic*
	(1)	(2)
Expressive Alignment 2016	−1.214*	1.610*
	(0.279)	(0.304)
Party: Democrats	0.974*	−0.186
	(0.178)	(0.227)
Party: Republican	−0.377	0.677*
	(0.219)	(0.174)
Ideology (1-7)	−0.096	0.243*
	(0.050)	(0.054)
Feeling Therm: Trump - Clinton	−0.027*	0.031*
	(0.002)	(0.002)
Racial Resentment	−0.188*	0.037
	(0.068)	(0.076)
Hostile Sexism	−0.045	0.001
	(0.073)	(0.078)
Authoritarianism	−0.098	−0.062
	(0.052)	(0.056)
Education	0.091*	0.070*
	(0.029)	(0.031)
Age (yrs)	0.073*	0.113*
	(0.018)	(0.019)
Female	−0.186	0.263*
	(0.122)	(0.125)
Race: Black	0.609	0.882
	(0.335)	(0.528)
Race: Hispanic	−0.105	0.555
	(0.328)	(0.445)
Race: Native American	0.850	−0.969
	(0.689)	(1.213)
Race: Other	0.405	0.634
	(0.401)	(0.480)
Race: White	0.231	0.845*
	(0.297)	(0.374)
Income	0.036*	0.026*
	(0.008)	(0.008)
Political Attention	0.890*	0.023
	(0.228)	(0.240)
Mode: Web	−0.403*	−0.179
	(0.130)	(0.132)
Expressive Alignment 2016 × Dem	0.914*	0.721
	(0.343)	(0.582)
Expressive Alignment 2016 × Rep	−0.666	−1.431*
	(0.538)	(0.370)
Constant	−2.726*	−5.610*
	(0.610)	(0.717)
Observations	3,289	3,289
Log Likelihood	-986.604	-919.035
Akaike Inf. Crit.	2,021.207	1,886.070
Note:		$^{*}p < 0.05$

Table A1.12: Candidate Choice in 2016 vs Expressive Alignment 2016

A1.7 Study 1b: Candidate Choice in 2020 vs Expressive Alignment 2016

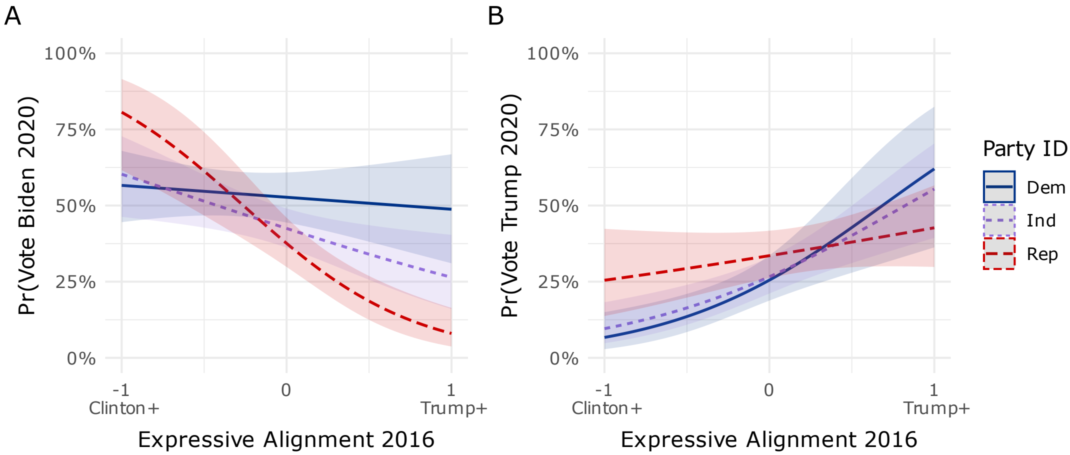

Figure A1.4. Marginal effects of Expressive Alignment in 2016 on candidate preference in 2020, by party identification (see Table A1.13).

	Dependent variable:	
	Vote Biden 2020	Vote Trump 2020
	(1)	(2)
Expressive Alignment 2016	−0.720*	1.231*
	(0.275)	(0.321)
Party: Democrats	0.411*	−0.058
	(0.183)	(0.224)
Party: Republican	−0.204	0.330
	(0.192)	(0.189)
Feeling Therm: Trump - Clinton	−0.018*	0.020*
	(0.002)	(0.002)
Ideology (1-7)	−0.132*	0.229*
	(0.054)	(0.061)
Racial Resentment	−0.282*	0.308*
	(0.072)	(0.082)
Hostile Sexism	−0.144	0.0003
	(0.079)	(0.086)
Authoritarianism	−0.057	0.041
	(0.056)	(0.061)
Education	0.041	0.137*
	(0.031)	(0.035)
Age (yrs)	0.070*	0.024
	(0.020)	(0.021)
Female	−0.019	0.128
	(0.130)	(0.137)
Race: Black	0.192	−0.528
	(0.378)	(0.571)
Race: Hispanic	0.077	−0.482
	(0.373)	(0.455)
Race: Native American	0.777	1.109
	(1.018)	(1.148)
Race: Other	0.272	0.111
	(0.458)	(0.524)
Race: White	0.277	0.156
	(0.330)	(0.372)
Income	0.028*	0.013
	(0.009)	(0.009)
Political Attention	0.245	−0.335
	(0.252)	(0.274)
Mode: Web	0.110	0.111
	(0.138)	(0.149)
Expressive Alignment 2016 × Dem	0.562	0.332
	(0.339)	(0.509)
Expressive Alignment 2016 × Rep	−1.214*	−0.841*
	(0.456)	(0.382)
Constant	−0.443	−5.265*
	(0.658)	(0.775)
Observations	2,281	2,281
Log Likelihood	-874.047	-763.773
Akaike Inf. Crit.	1,796.095	1,575.547
Note:		*$p < 0.05$

Table A1.13: Candidate Choice in 2020 vs Expressive Alignment 2016

A1.8 Study 1b: Candidate Choice in 2024 vs Expressive Alignment 2016

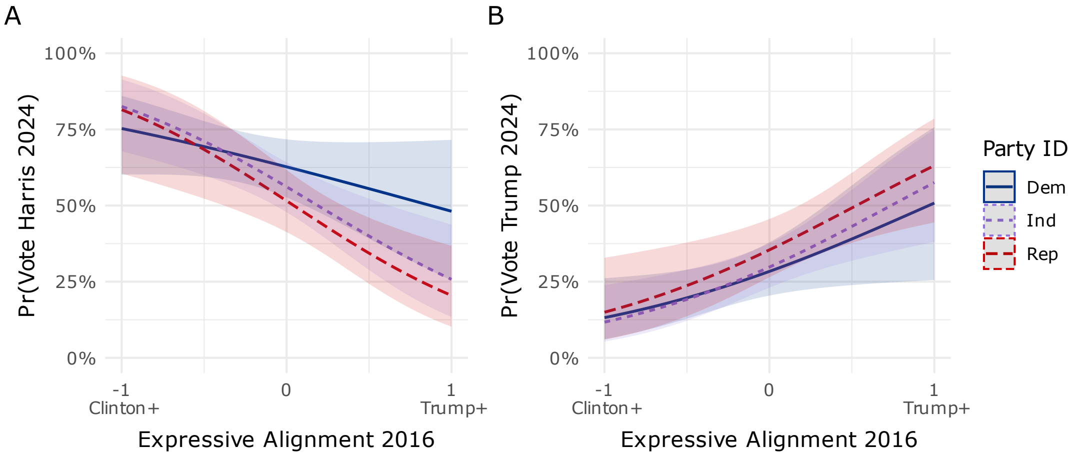

Figure A1.5. Marginal effects of Expressive Alignment in 2016 on candidate preference in 2024, by party identification (see Table A1.14).

	Dependent variable:	
	Vote Harris 2024	Vote Trump 2024
	(1)	(2)
Expressive Alignment 2016	−1.306*	1.162*
	(0.375)	(0.381)
Party: Democrats	0.272	−0.069
	(0.228)	(0.239)
Party: Republican	−0.185	0.257
	(0.223)	(0.223)
Feeling Therm: Trump - Clinton	−0.023*	0.025*
	(0.003)	(0.003)
Ideology (1-7)	−0.192*	0.275*
	(0.073)	(0.076)
Racial Resentment	−0.329*	0.262*
	(0.095)	(0.099)
Hostile Sexism	−0.238*	0.083
	(0.104)	(0.106)
Authoritarianism	−0.090	0.110
	(0.072)	(0.073)
Education	0.076	−0.037
	(0.041)	(0.042)
Age (yrs)	0.103*	−0.103*
	(0.027)	(0.027)
Female	−0.087	−0.029
	(0.167)	(0.168)
Race: Black	0.745	−0.997
	(0.513)	(0.557)
Race: Hispanic	0.555	−0.391
	(0.491)	(0.510)
Race: Native American	−0.290	−0.484
	(1.318)	(1.251)
Race: Other	0.550	−0.568
	(0.600)	(0.618)
Race: White	0.665	−0.743
	(0.425)	(0.437)
Income	0.017	−0.014
	(0.012)	(0.012)
Political Attention	0.370	−0.283
	(0.332)	(0.338)
Mode: Web	−0.115	0.059
	(0.182)	(0.185)
Expressive Alignment 2016 × Dem	0.711	−0.203
	(0.485)	(0.537)
Expressive Alignment 2016 × Rep	−0.113	−0.025
	(0.518)	(0.500)
Constant	0.191	−0.921
	(0.879)	(0.907)
Observations	1,782	1,782
Log Likelihood	-545.267	-527.329
Akaike Inf. Crit.	1,138.535	1,102.658
Note:		*$p < 0.05$

Table A1.14: Candidate Choice in 2024 vs Expressive Alignment 2016

A2. Study 2: Validated Turnout vs Expressive Engagement

A2.1 Number of character summary statistics for Most Important Problem response by survey mode and year

	Max		Mean		Median	
Problem #	FTF	Web	FTF	Web	FTF	Web
Problem 1	624	739	66.6	32.5	42	24
Problem 2	980	816	59.8	28.9	30	21
Problem 3	450	768	44.4	23.6	12	18
Problem 4	969	440	19.8	16.3	10	12

Note: Face-to-face: n = 1,180, Web: n = 3,090.

Table A2.15: Number of character summary statistics for Most Important Problem responses by survey mode16 (2016)

	Max			Mean			Median		
Problem #	Tele	Video	Web	Tele	Video	Web	Tele	Video	Web
Problem 1	906	593	1198	83.8	63.4	69.2	30	22	41
Problem 2	628	584	1197	65.2	57.7	45.9	20	18	24
Problem 3	542	888	1194	48.9	53.2	28.9	9	12	0
Problem 4	186	690	958	19.1	21.0	23.5	12	11	10

Note: Tele: n = 139, Video: n = 359, Web: n = 7,782.

Table A2.16: Number of character summary statistics for Most Important Problem responses by survey mode16 (2020)

I do not present summary statistics for 'Most Important Problem' responses by survey mode in 2024 as the validated turnout data is not yet available and, therefore, calculating the Expressive Engagement Scale in 2024 has no application.

A2.2 Study 2: Interpreting Expressive Engagement

In the ANES 2016 Time Series Study, subjects were asked three times about the most important problems in country and then a fourth time about which of the prior three is the most important problem. As discussed in Equation 5, I combine these four responses into a measure I call *Expressive Engagement* to reflect how this scale captures something similar to other forms of civic and political engagement but through private articulation about issues rather than public action.

Conceptually, this measure draws on the text as behavior idea that more verbose responses indicate greater cognitive and affective effort, and willingness to articulate political concerns. To better understand what is being captured, I use simple correlations and Principal Component Analysis to show how this variable is related to other comparable concepts and measures. Table A2.17 shows that Expressive Engagement is modestly and positively correlated with political knowledge ($r = 0.23$), civic and behavioral participation ($r \approx 0.18$–0.20), political interest ($r = 0.16$), and political discussing politics with others ($r = 0.15$).

Figure A2.6 shows a Principal Component Analysis (PCA) of Expressive Engagement with those same variables. The PCA results show the Expressive Engagement variable, `nchar_problems_ihs`, loads most heavily on the second principal component (Dim2 / PC 2), which distinguishes articulation (e.g., verbosity) from more concrete political behaviors (e.g., donating, volunteering). It also loads, to a lesser degree, on the first principal component (Dim1 / PC 1), which reflects conventional political engagement, here defined by political interest and knowledge (see Section A.8). These results suggest that Expressive Engagement captures both traditional forms of participation and a distinct, more symbolic mode of political involvement.

Term	Correlation
Political Knowledge Scale	0.227
Civic Engagement Scale	0.196
Political Engagement Scale	0.181
Political Interest	0.164
discuss_pol_days	0.151
Likely Vote	0.147

Table A2.17: Select Correlates of Expressive Engagement

I also note that predicting validated voter turnout is particularly difficult when the model includes a self-reported measure of likely vote intention that dominates most turnout models. This variable typically absorbs a substantial share of the explained variance, often leaving little room for additional predictors to contribute meaningfully. The fact that Expressive Engagement remains a significant predictor of validated turnout, even when controlling for likelihood to vote, suggests that it captures something distinct from self-reported intent. The likelihood to vote measure plausibly reflects a conscious commitment to voting. In contrast, the Expressive Engagement seems to capture some additional behavioral indicator of motivation to participate, revealed not through self-report but through the effort of articulating political concerns. This distinction underscores the value of open-ended response data as a complementary and behavioral measure of political engagement.

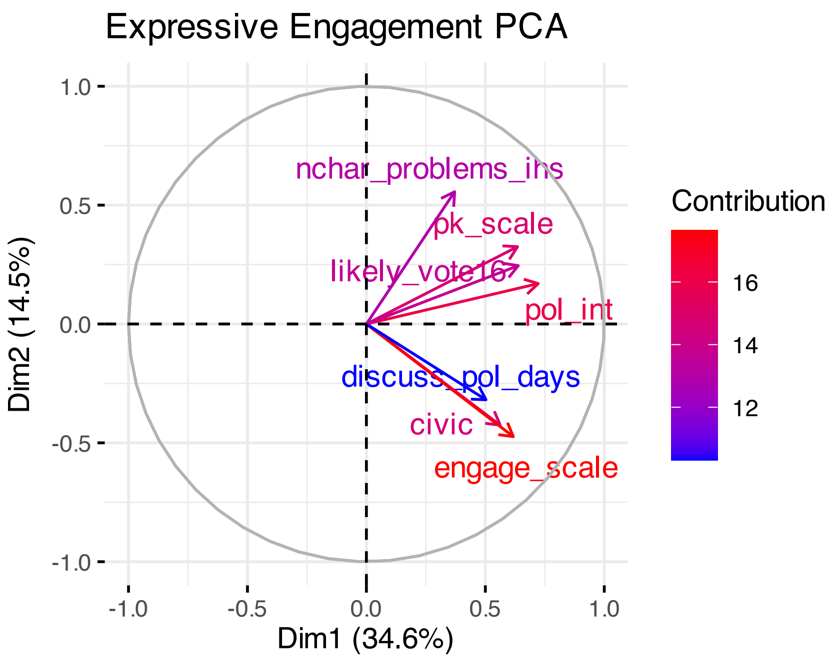

Figure A2.6. Principal Component Analysis suggests that Expressive Engagement (nchar_problems_ihs) loads more on the second principal component, which appears to distinguish expressive articulation (e.g., verbosity) from behavioral political action (e.g., volunteering or donating). The first component, capturing a general political engagement axis—from low (left) to high (right)—is defined primarily by political interest and knowledge. These results suggest that Expressive Engagement reflects both conventional forms of participation and a distinct, more symbolic mode of political involvement.

A2.3 Study 2: Validated Turnout in 2016 and 2020 vs Expressive Engagement in 2016

	Dependent variable:	
	Turnout 2016	Turnout 2020
	glm: quasibinomial *link = logit*	*glm: quasibinomial* *link = logit*
	(1)	(2)
Expressive Engagement 2016	0.240*	0.247*
	(0.049)	(0.095)
Party: Democrats	0.046	0.066
	(0.111)	(0.129)
Party: Republican	0.071	0.371*
	(0.113)	(0.137)
Female	0.104	−0.023
	(0.088)	(0.106)
Education	0.047*	0.093*
	(0.021)	(0.026)
Age (yrs)	0.120*	0.110*
	(0.013)	(0.016)
Race: Hispanic	0.185	−0.129
	(0.183)	(0.228)
Race: Other	0.067	−0.089
	(0.197)	(0.247)
Race: White	0.345*	0.281
	(0.145)	(0.178)
Income	0.036*	0.025*
	(0.006)	(0.007)
Political Attention	−0.459*	−0.122
	(0.178)	(0.218)
Mode: Web	−0.081	−0.137
	(0.098)	(0.121)
Likely Vote	0.605*	0.264*
	(0.046)	(0.057)
Constant	−4.050*	−3.054*
	(0.325)	(0.415)
Observations	3,240	2,210
Note:		* $p < 0.05$

Table A2.18: Validated Turnout incorporating match probability in 2016 and 2020 vs Expressive Engagement in 2016

A2.4 Study 2: Validated Turnout in 2020 vs Expressive Engagement in 2020

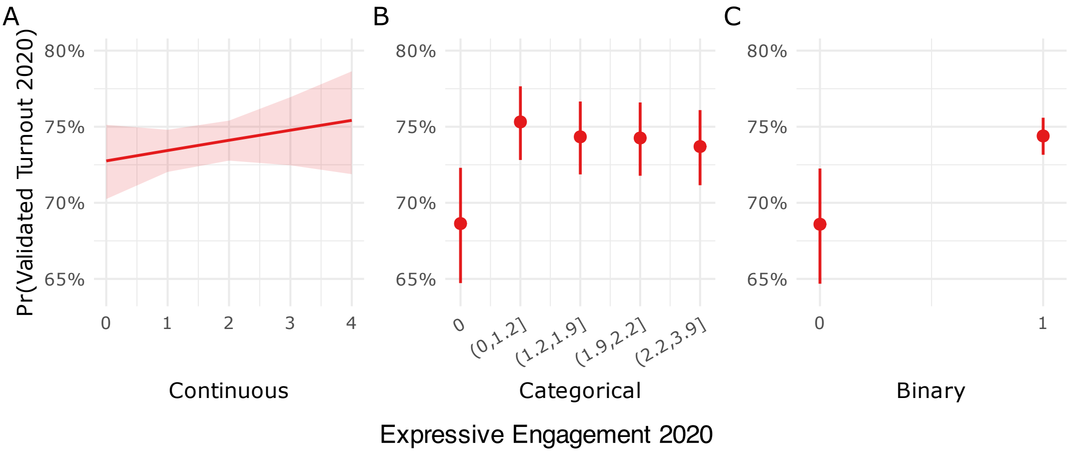

Figure A2.7. Marginal effects of Expressive Engagement in 2020 on predicted probability of validated turnout in 2020. Panel A uses Model 3 from Table A2.21, Panel B uses Model 3 from Table A2.22 and Panel C uses Model 3 from Table A2.23.

While Expressive Engagement in 2016 is a strong predictor of validated turnout in 2016 and 2020, the Expressive Engagement measure calculated from the 2020 ANES is a much weaker predictor of turnout in 2020 (see Table A2.21). To better understand what might be driving differences across the two cycles, I compare three specifications of Expressive Engagement in 2020—continuous, categorical and binary—and also compare the distributions of Expressive Engagement by year and mode.

Figure A2.7 shows the marginal effect of three different specifications. While all three plots show a positive association between Expressive Engagement in 2020 and validated turnout in 2020, only the categorical and binary specifications are statistically significant predictors of turnout in models that include self-reported likelihood to vote (see Model 3 in Tables A2.22 and A2.23, respectively). One insight suggested by Figure A2.7 is that the functional form of the relationship is not monotonically increasing across the observed range of responses increasing. In 2020, the main distinction appears to be between non-responders and any-responders. In both the categorical and binary specifications, the marginal effect of any-response is an increase of approximately five to six percentage points in the predicted probability of turning out to vote.

Equation 9 shows the function for dichotomizing Expressive Engagement. Equation 10 shows the function used to transform the continuous variable into a categorical specification. In short, Equation 10 keeps all zero responses as zero and then divides the remaining non-zero responses into quartiles (where $Q_j (j \in 1, 2, 3, 4)$ denotes the quartile of Expressive Engagement among non-zero responders).

$$\text{Expressive Engagement (Binary)} = \begin{cases} 0 & \text{if Expressive Engagement (Continuous)} = 0 \\ 1 & \text{if Expressive Engagement (Continuous)} > 0 \end{cases} \tag{9}$$

$$\text{Expressive Engagement (Categorical)} = \begin{cases} 0 & \text{if Expressive Engagement (Continuous)} = 0 \\ Q_j & \text{if Expressive Engagement (Continuous)} > 0 \end{cases} \tag{10}$$

To further understand these result, in Figure A2.8 I plot five histograms that compare Expressive Engagement (e.g., the transformed distribution of responses to the "most important problems" questions) in ANES 2016 and 2020 by survey mode. Though the overall patterns are broadly similar, there is a notable spike between 0 and 1 among those surveyed in 2020 via the web.

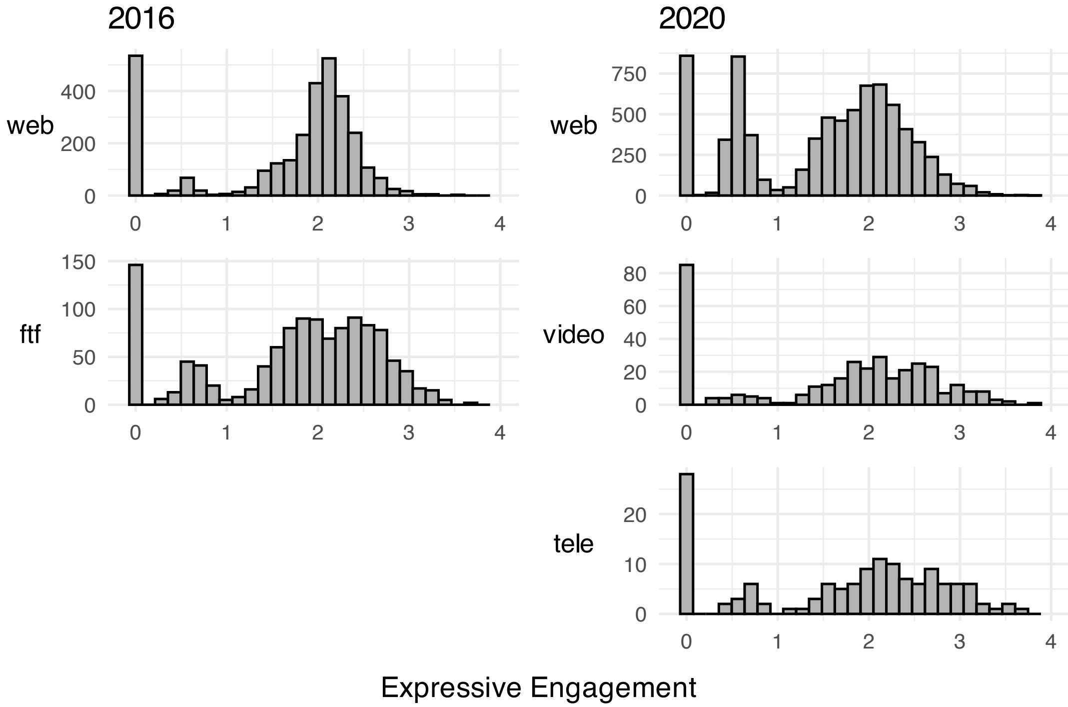

Figure A2.8. Distribution of Expressive Engagement across modes in 2016 and 2020. Each histogram represents the distribution of IHS-normalized total character count across four open-ended 'most important problem' responses. The spikes at zero reflect nonresponse to all items; in 2020, a secondary peak just below 1 captures minimal engagement (typically one short answer); and the broader distribution represents fuller engagement. Note: As focus is on comparing distributions, the y-axes vary. In 2016, web N = 3090, face-to-face n = 1180. In 2020, web n = 7,782, video n = 359, tele n = 139.

	Face-to-Face		Web	
# MIP Answered	n	Percent	n	Percent
0	146	12%	535	17%
1	120	10%	113	4%
2	10	1%	14	0%
3	179	15%	319	10%
4	725	61%	2109	68%

Table A2.19: ANES 2016 'Most Important Problems' Answered

Tables A2.19 and A2.20 show how many most important problem questions were answered in both waves across all modes. Looking at the bottom rows of both tables, we can see that the modal response pattern was to answer all four prompts. A key difference emerges between years and mode, however, in the proportion of respondents who answered only the first question: just 233 individuals of 4,270 (5.5%) in 2016 gave only one response compared to 1,704 individuals of 8,280 (20.6%) in 2020. This sharp, 3.8 times increase in drop-off is concentrated among 2020 web respondents and could, plausibly, reflect something like fatigue.

Due to COVID, 2020 was predominantly web-based (94%), while 2016 included both web (72%) and face-to-face interviews (28%). The 2020 web respondents appear more prone to partial

	Tele		Video		Web	
# MIP Answered	n	Percent	n	Percent	n	Percent
0	28	20%	85	24%	859	11%
1	13	9%	24	7%	1667	21%
2	0	0%	0	0%	62	1%
3	23	17%	43	12%	1806	23%
4	75	54%	207	58%	3388	44%

Table A2.20: ANES 2020 'Most Important Problems' Answered

completion, especially stopping after the first item. Other partial response combinations, such as answering exactly two questions, remained rare in both years, while answering exactly three questions was relatively common—particularly in 2020. Importantly, however, once responses are aggregated and transformed into the Expressive Engagement scale (Figure A2.8), these multi-item completions collapse into a single upper mode. At that level, the 2020 web data exhibit an approximately trimodal pattern: a mass at zero (non-responders), a secondary spike reflecting minimal engagement (typically one short answer), and a broader distribution corresponding to fuller engagement.

As these models do not rely on the panel structure of the ANES, the models include all 2020 subjects, not just those who completed the 2016 ANES, and all covariates are from 2020 as well.

Though the categorical and binary specifications of Expressive Engagement in 2020 suggest that the scale continues to be meaningful, additional work is necessary to determine whether the observed differences between 2016 and 2020 reflect features of the scale's construction or idiosyncratic aspects of the 2020 election (e.g., COVID-related disruptions, sampling methodology, or survey design). Another related possibility is that, due to COVID, the 2020 election saw a dramatic shift to increased voting by mail. This increase in ease of voting may be reflected in the ANES data. In 2016, about 14 percent of ANES subjects who were not registered ultimately voted. In 2020, that number more than doubles to 36 percent (see Section A1.3). If voting becomes *less* effortful, then it is possible a measure of *more* effortful political engagement could become less predictive of turnout.

	Dependent variable:		
	Validated Turnout 2020		
	glm: quasibinomial *link = logit*		
	(1)	(2)	(3)
Expressive Engagement 2020	0.077*	0.057	0.035
	(0.033)	(0.034)	(0.036)
Party: Democrats	0.185*	0.160*	0.074
	(0.072)	(0.072)	(0.076)
Party: Republican	0.245*	0.231*	0.161*
	(0.076)	(0.076)	(0.079)
Female	0.070	0.107	0.070
	(0.059)	(0.059)	(0.062)
Education	0.090*	0.083*	0.065*
	(0.016)	(0.016)	(0.017)
Age (yrs)	0.024*	0.022*	0.020*
	(0.002)	(0.002)	(0.002)
Race: Black	0.232	0.201	0.057
	(0.172)	(0.173)	(0.180)
Race: Hispanic	0.080	0.046	−0.061
	(0.171)	(0.171)	(0.179)
Race: Other	0.328	0.279	0.142
	(0.186)	(0.187)	(0.194)
Race: White	0.639*	0.603*	0.466*
	(0.150)	(0.150)	(0.156)
Income	0.035*	0.034*	0.027*
	(0.005)	(0.005)	(0.005)
Mode: Video	0.205	0.194	0.117
	(0.262)	(0.262)	(0.332)
Mode: Web	0.001	−0.003	−0.074
	(0.223)	(0.223)	(0.301)
Political Attention		−0.125*	0.003
		(0.030)	(0.033)
Likely Vote			0.530*
			(0.039)
Constant	−1.892*	−1.418*	−3.570*
	(0.296)	(0.317)	(0.414)
Observations	6,598	6,597	6,236
Note:			*$p < 0.05$

Table A2.21: Validated turnout incorporating match probability in 2020 vs Expressive Engagement 2020 (Continuous) using quasibinomial models

	Dependent variable:		
	Validated Turnout 2020		
	glm: quasibinomial link = logit		
	(1)	(2)	(3)
Expressive Engagement 2020: Q1	0.366*	0.362*	0.332*
	(0.105)	(0.105)	(0.110)
Expressive Engagement 2020: Q2	0.334*	0.317*	0.280*
	(0.104)	(0.104)	(0.110)
Expressive Engagement 2020: Q3	0.368*	0.332*	0.276*
	(0.105)	(0.105)	(0.111)
Expressive Engagement 2020: Q4	0.361*	0.315*	0.247*
	(0.105)	(0.106)	(0.111)
Party: Democrats	0.183*	0.159*	0.073
	(0.072)	(0.072)	(0.076)
Party: Republican	0.237*	0.223*	0.152
	(0.076)	(0.076)	(0.079)
Female	0.068	0.105	0.068
	(0.059)	(0.059)	(0.062)
Education	0.090*	0.083*	0.065*
	(0.016)	(0.016)	(0.017)
Age (yrs)	0.024*	0.023*	0.020*
	(0.002)	(0.002)	(0.002)
Race: Black	0.227	0.195	0.053
	(0.172)	(0.173)	(0.180)
Race: Hispanic	0.087	0.052	−0.053
	(0.171)	(0.172)	(0.179)
Race: Other	0.332	0.282	0.146
	(0.186)	(0.187)	(0.195)
Race: White	0.636*	0.598*	0.465*
	(0.150)	(0.150)	(0.156)
Income	0.035*	0.034*	0.027*
	(0.005)	(0.005)	(0.005)
Mode: Video	0.209	0.198	0.118
	(0.262)	(0.263)	(0.333)
Mode: Web	−0.055	−0.060	−0.130
	(0.224)	(0.225)	(0.303)
Political Attention		−0.127*	0.002
		(0.030)	(0.033)
Likely Vote			0.532*
			(0.039)
Constant	−2.056*	−1.583*	−3.742*
	(0.301)	(0.322)	(0.419)
Observations	6,598	6,597	6,236
Note:			*$p < 0.05$

Table A2.22: Validated turnout incorporating match probability in 2020 vs Expressive Engagement 2020 (Categorical, reference category is zero characters) using quasibinomial models.

	Dependent variable:		
	Validated Turnout 2020		
	glm: quasibinomial *link = logit*		
	(1)	(2)	(3)
Expressive Engagement 2020 (Binary)	0.357*	0.332*	0.285*
	(0.090)	(0.090)	(0.095)
Party: Democrats	0.183*	0.158*	0.073
	(0.072)	(0.072)	(0.076)
Party: Republican	0.237*	0.225*	0.157*
	(0.076)	(0.076)	(0.079)
Female	0.069	0.106	0.068
	(0.059)	(0.059)	(0.062)
Education	0.090*	0.082*	0.064*
	(0.016)	(0.016)	(0.016)
Age (yrs)	0.025*	0.023*	0.020*
	(0.002)	(0.002)	(0.002)
Race: Black	0.227	0.195	0.054
	(0.172)	(0.173)	(0.180)
Race: Hispanic	0.087	0.051	−0.056
	(0.171)	(0.171)	(0.179)
Race: Other	0.332	0.280	0.143
	(0.186)	(0.187)	(0.194)
Race: White	0.635*	0.597*	0.463*
	(0.150)	(0.150)	(0.156)
Income	0.035*	0.034*	0.027*
	(0.005)	(0.005)	(0.005)
Mode: Video	0.208	0.199	0.118
	(0.262)	(0.263)	(0.333)
Mode: Web	−0.056	−0.054	−0.121
	(0.223)	(0.224)	(0.303)
Political Attention		−0.125*	0.005
		(0.030)	(0.033)
Likely Vote			0.531*
			(0.039)
Constant	−2.055*	−1.582*	−3.733*
	(0.300)	(0.321)	(0.419)
Observations	6,598	6,597	6,236
Note:			*$p < 0.05$

Table A2.23: Validated turnout incorporating match probability in 2020 vs Expressive Engagement 2020 (Binary) using quasibinomial models.

A2.5 Study 2: Validated Turnout vs Expressive Alignment

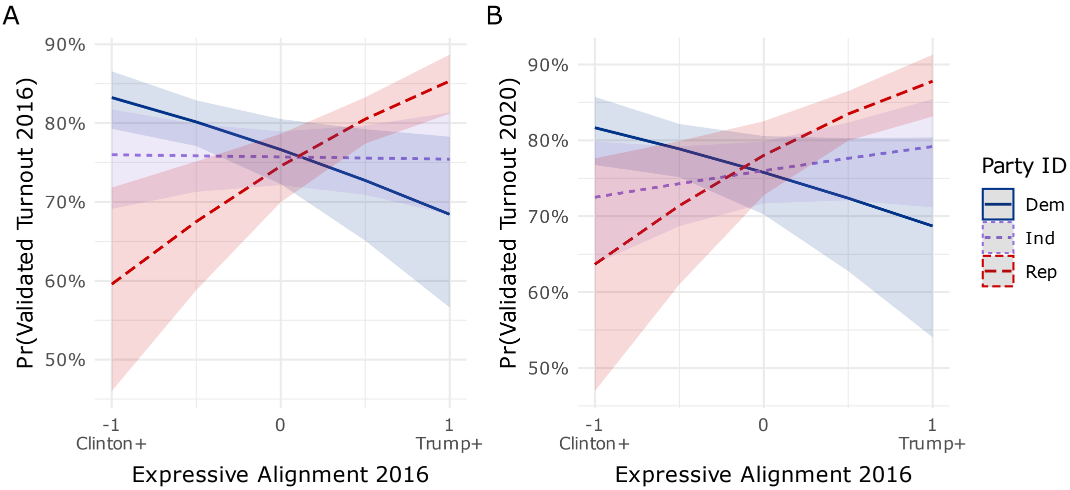

Figure A2.9. Marginal effects of Expressive Alignment in 2016 on validated turnout in 2016 and 2020 with logistic regression model (see Table A2.24).

As a secondary analysis, I estimate the probability of turning out to vote in 2016 and 2020 as a function of Expressive Alignment measured in 2016. Figure A2.9 presents marginal effects from models that exclude the Trump-Clinton feeling thermometer difference to better isolate the relationship between open-ended candidate evaluations and turnout. Across both panels, Expressive Alignment is positively associated with turnout among Republicans and, to a lesser extent, Independents, but negatively associated with turnout among Democrats.

In Panel A, the predicted probability of validated turnout in 2016 rises from 60 to 85 percent among Republicans as Expressive Alignment increases from –1 to +1. Among Independents, the slope is essentially flat, rising only slightly from 76 to 75 percent. Among Democrats, Expressive Alignment is negatively associated with turnout, declining from 83 to 68 percent.

A similar pattern appears in Panel B for the 2020 election: turnout increases from 64 to 88 percent among Republicans, from 72 to 79 percent among Independents, and decreases from 82 to 69 percent among Democrats. While the signs and magnitudes of the estimated associations are consistent across years, the estimates in 2020 carry more uncertainty—particularly for Democrats and Independents—and interaction terms are not statistically significant.

Overall, these results show that partisan-congruent expressive behavior strongly predicts turnout: individuals whose open-ended evaluations align with their party identity are more likely to vote. However, candidate evaluations alone are insufficient; turnout reflects the interaction between partisan identity and expressive alignment rather than either factor in isolation.In 2016, the pattern is especially clear: those exhibiting partisan-congruent expressive behavior are more likely to vote (e.g., Democrats expressing more in the "for Clinton" and "against Trump" prompts, Republicans expressing more in the "for Trump" and "against Clinton" prompts). By contrast, those whose expressive behavior is partisan-discordant show substantially lower turnout, suggesting that expressive ambivalence may lead to demobilization rather than defection.

To assess whether Expressive Alignment adds predictive power beyond conventional measures of candidate-based affective polarization, I also estimate parallel models that include the difference in feeling thermometer scores for Trump and Clinton (Trump minus Clinton). These results are presented in Table A2.24. Including the thermometer difference slightly attenuates the interaction between Expressive Alignment and Democratic identity in 2016, rendering it non-significant. However, the interaction with Republican identity remains large and statistically significant. This suggests that while some of the turnout effect among Democrats may be explained by more conventional measures of affective polarization, Expressive Alignment still captures unique behavioral information, especially among Republicans.

Finally, I do not include self-reported likelihood of voting (which is used in models predicting Expressive Engagement) because I treat attitudes about turnout as downstream of the affective and cognitive processes captured by Expressive Alignment (see Section A1.6). Including likely vote as a control would risk conditioning on a post-treatment variable and thus obscure the key relationships of interest. Models with and without the candidate affect gap can be seen in Table A2.24. When the feeling thermometer control is included, the joint effect of Expressive Alignment and party identification remains significant for Republicans in 2016 but not in 2020, and is not significantly different than the nearly flat Independent's slope in either cycle for Democrats.

	Dependent variable:			
	Turnout 2016		Turnout 2020	
	glm: quasibinomial *link = logit*		*glm: quasibinomial* *link = logit*	
	(1)	(2)	(3)	(4)
Expressive Alignment 2016	−0.015	0.183	0.309	0.309
	(0.150)	(0.181)	(0.228)	(0.228)
Party: Democrats	0.051	−0.013	−0.037	−0.037
	(0.127)	(0.156)	(0.163)	(0.163)
Party: Republican	−0.063	0.114	0.152	0.152
	(0.127)	(0.159)	(0.163)	(0.163)
Feeling Therm: Trump - Clinton			−0.001	−0.001
			(0.002)	(0.002)
Female	0.140	0.011	0.027	0.027
	(0.084)	(0.104)	(0.105)	(0.105)
Education	0.077*	0.107*	0.104*	0.104*
	(0.020)	(0.025)	(0.025)	(0.025)
Age (yrs)	0.134*	0.121*	0.123*	0.123*
	(0.013)	(0.016)	(0.016)	(0.016)
Race: Hispanic	0.132	−0.091	−0.100	−0.100
	(0.172)	(0.220)	(0.222)	(0.222)
Race: Other	0.103	0.012	0.001	0.001
	(0.190)	(0.244)	(0.245)	(0.245)
Race: White	0.412*	0.375*	0.370*	0.370*
	(0.140)	(0.175)	(0.177)	(0.177)
Income	0.036*	0.025*	0.027*	0.027*
	(0.006)	(0.007)	(0.007)	(0.007)
Political Attention	0.132	0.043	0.060	0.060
	(0.162)	(0.206)	(0.207)	(0.207)
Mode: Web	−0.192*	−0.158	−0.155	−0.155
	(0.093)	(0.117)	(0.117)	(0.117)
Expressive Alignment 2016 × Dem	−0.399	−0.537*	−0.537	−0.537
	(0.224)	(0.271)	(0.274)	(0.274)
Expressive Alignment 2016 × Rep	0.702*	0.523	0.475	0.475
	(0.242)	(0.300)	(0.303)	(0.303)
Constant	−1.840*	−1.812*	−1.823*	−1.823*
	(0.267)	(0.344)	(0.347)	(0.347)
Feeling Thermometer?	No	No	Yes	Yes
Observations	3,380	2,300	2,278	2,278
Note:				*$p < 0.05$

Table A2.24: Predicted probability of validated turnout incorporating match probability in 2016 and 2020 vs Expressive Alignment in 2016.

A3. Study 3: Text as Multilingual Instrument Validation

Table A3.25 presents basic summary statistics of the number of respondents by language. The Afrobarometer asks subjects a democracy battery made up of ten questions about whether liberal democracy is important. The survey also asked subjects three times, "What, if anything, does 'democracy' mean to you?" Interviewers were instructed to, "Read the question in the language of the interview, but always state the word 'democracy' in English. Only translate 'democracy' into local language if respondent does not understand the term in the official national language. Record whether respondent understood word in English or required a local language translation. Be sure to ask ALL questions of ALL respondents, even if they have difficulty understanding the term 'democracy'" (Afrobarometer 2014).

A3.1 Study 3: Table for Afrobarometer Summary Statistics

Language	n	Percent
English	34,838	65%
French	14,116	26%
Portuguese	4,707	9%
NA	274	1%
Total	53,935	-

Table A3.25: Frequency table of Languages within Afrobarometer

A3.2 Study 3: Table for Afrobarometer Regression Results

	Dependent variable:			
	Number of Characters: 'What Democracy Means to You?'			
	English	French	Portuguese	Interaction
	(1)	(2)	(3)	(4)
Importance of Democracy	0.006*	0.010*	0.011*	0.006*
	(0.001)	(0.001)	(0.002)	(0.001)
understood_demLocal language	−0.041*	−0.112*	0.204*	
	(0.010)	(0.015)	(0.061)	
Lang: French				−0.078
				(0.042)
Lang: Portuguese				−0.224*
				(0.058)
gender	−0.050*	−0.091*	−0.015	−0.064*
	(0.008)	(0.011)	(0.027)	(0.006)
Education	0.053*	0.026*	0.041*	0.046*
	(0.002)	(0.003)	(0.008)	(0.002)
Age (yrs)	−0.0002	0.001*	0.001	0.0004
	(0.0003)	(0.0004)	(0.001)	(0.0002)
Income	0.001	−0.005*	0.007*	−0.001
	(0.001)	(0.001)	(0.003)	(0.001)
race5Arab-North African	−0.111*	0.172		−0.109*
	(0.012)	(0.224)		(0.011)
race5Colored-Mixed Race	−0.043	0.089	−0.097*	−0.078*
	(0.030)	(0.148)	(0.040)	(0.022)
Race: Other	−0.011	0.121	−0.149	0.004
	(0.023)	(0.148)	(0.160)	(0.023)
Race: Asian	−0.057*	−0.022	0.345	−0.052*
	(0.022)	(0.243)	(0.227)	(0.021)
dem_importance:langFrench				0.004*
				(0.001)
dem_importance:langPortuguese				0.002
				(0.002)
Constant	3.398*	3.444*	2.950*	3.416*
	(0.030)	(0.045)	(0.099)	(0.027)
Observations	27,459	12,279	3,111	42,849
Log Likelihood	-124,937.600	-55,455.710	-13,957.880	-194,620.200
Akaike Inf. Crit.	249,897.100	110,933.400	27,935.760	389,268.400
Note:				*p < 0.05

Table A3.26: Number of characters about democracy means vs democracy battery interacted with language

A3.3 Study 3: Afrobarometer Questions

- Q30: Question: Which of these three statements is closest to your own opinion?
- Q30. Support for democracy
  - Statement 1: Democracy is preferable to any other kind of government.
  - Statement 2: In some circumstances, a non-democratic government can be preferable.
  - Statement 3: For someone like me, it doesn't matter what kind of government we have.
  - Response options: 1=Statement 3: Doesn't matter, 2=Statement 2: Sometimes non-democratic preferable, 3=Statement 1: Democracy preferable, 9=Don't know, 98=Refused to answer, -1=Missing
- Q31-Q39:
  - Question prompt: Which of the following statements is closest to your view? Choose Statement 1 or Statement 2.
  - Response options: 1=Agree very strongly with Statement 1, 2=Agree with Statement 1, 3=Agree with Statement 2, 4=Agree very strongly with Statement 2, 5=Agree with neither, 9=Don't know, 98=Refused to answer, -1=Missing
- Q31. Government gets things done but no citizen influence vs. government accountable to citizens
  - Statement 1: It is more important to have a government that can get things done, even if we have no influence over what it does.
  - Statement 2: It is more important for citizens to be able to hold government accountable, even if that means it makes decisions more slowly.
- Q32. Choose leaders through elections vs. other methods
  - Statement 1: We should choose our leaders in this country through regular, open and honest elections.
  - Statement 2: Since elections sometimes produce bad results, we should adopt other methods for choosing this country's leaders.
- Q33. Political parties divisive vs. many parties needed
  - Statement 1: Political parties create division and confusion; it is therefore unnecessary to have many political parties in [ENTER COUNTRY].
  - Statement 2: Many political parties are needed to make sure that [ENTER NATIONALITY] have real choices in who governs them.
- Q34. President monitored by parliament vs. free to act on own
  - Statement 1: Parliament should ensure that the President explains to it on a regular basis how his government spends taxpayers' money.
  - Statement 2: The President should be able to devote his full attention to developing the country rather than wasting time justifying his actions.
- Q35. Opposition parties examine government vs. cooperate
  - Statement 1: After losing an election, opposition parties should monitor and criticize the government in order to hold it accountable.
  - Statement 2: Once an election is over, opposition parties and politicians should accept defeat and cooperate with government to help it develop the country.
- Q36. Media checks government vs. avoid negative reporting

  - Statement 1: The news media should constantly investigate and report on government mistakes and corruption.
  - Statement 2: Too much reporting on negative events, like government mistakes and corruption, only harms the country.

- Q37. Parliament makes laws vs. president does
  - Statement 1: Members of Parliament represent the people; therefore they should make laws for this country, even if the President does not agree.
  - Statement 2: Since the President represents all of us, he should pass laws without worrying about what Parliament thinks.

- Q38. President free to act vs. obey the laws and courts
  - Statement 1: Since the President was elected to lead the country, he should not be bound by laws or court decisions that he thinks are wrong.
  - Statement 2: The President must always obey the laws and the courts, even if he thinks they are wrong.

- Q39. Presidential two term limit vs. no term limits
  - Statement 1: The Constitution should limit the president to serving a maximum of two terms in office.
  - Statement 2: There should be no constitutional limit on how long the president can serve.

A3.4 Study 3: Importance of Democracy Scale: Sum vs. Mean

In the 10-question battery about the importance of democracy, each question is scaled or rescaled from 1 to 5. The scale is coded so that 1 is the least supportive of democratic institutions, while 5 is the most supportive. Subjects who respond, "Agree with neither" are coded at the center or 3. I take the sum of the 10 questions to create the *Importance of Democracy* Scale. I opt for the sum, rather than the mean for methodological and substantive reasons (the results are robust to either approach). As with most surveys, there is some missingness across the 10 questions. Specifically, responses coded as "9=Don't know, 98=Refused to answer, -1=Missing," were coded as missing. The level of missingness ranges from about 2 to 11 percent across the 10 questions in the battery. Figure A3.10 presents the raw data in a scatter plot of the number of characters in response to the "What democracy means to you?" prompts versus the Importance of Democracy Scale.

Building on the earlier results about informative nonresponse, I treat the missing values as meaningful and indicative of low political engagement. Put another way, the ten question battery captures both attitudes about democratic governance and, in the missingness, a degree of concern about politics at all. For the purposes of validating the 'Importance of Democracy' battery via a behavioral text measure, both dimensions of the answers are worthy of capture.

For example, imagine two subjects: one who responds to every question with the most pro-democracy answer (e.g., their vector of responses would be ten 5s), and a second subject who responds to one question with the most pro-democracy answer but responds "Don't know" to the other nine questions (e.g., one 5, and 9 NAs). If the scale of ten questions is calculated with a mean, assuming NAs are dropped, both subjects would be scaled to the same value of 5 (e.g., 50/10 = 5/1 = 5). In contrast, with a sum, the first subject would be scaled to 50 while the second subject would be coded to 5 (again, with NAs removed).

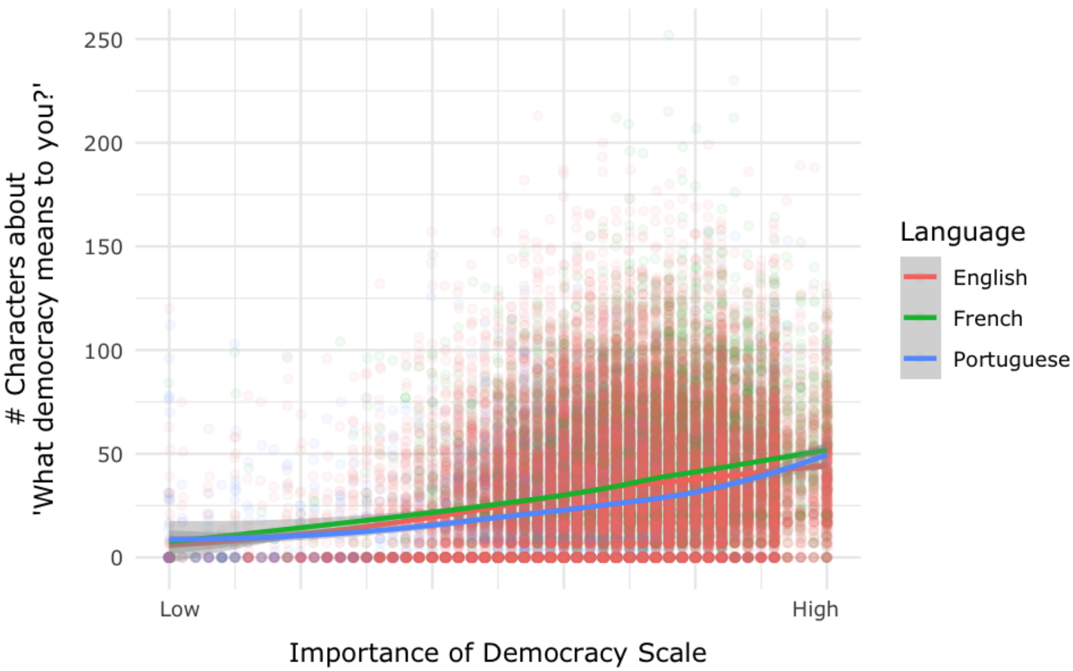

Figure A3.10. Scatter plot of raw data showing number of characters about 'What democracy means to you?' versus Importance of Democracy Scale with smoothed Loess curves by language.

A4. Study 5: Text as Manipulation Check of Social Exclusion

Can open-ended text prompts help reveal thoughts or feelings that are ambiguous or complex? Much of what people think or feel in a given moment is often opaque, even to themselves (Wilson 2004). In this section, I examine whether metadata about writing can help illuminate whether an experimental manipulation inducing subtle social exclusion was successful.

Kuo, Malhotra, and Mo (2017) (henceforth KMM) study how feelings of social exclusion might influence political attitudes, particularly among Asian Americans. As noted previously, KMM randomly induce feelings of social exclusion in treated white and Asian American subjects. Subjects then completed a survey and, across four different questions, wrote lists of things they liked and disliked about both the Democratic and Republican Parties. Results suggested that the joint effect of being Asian American and receiving a microaggression treatment, versus being white in the control condition, caused a negative shift in attitudes toward the Republican Party and more positive affect toward the Democratic Party.

As a form of manipulation check, KMM further asked subjects to "list as many US politicians [as] they could think of on the spot" (27). They note, "if the racial microaggression offends Asians, they may desire to compensate by showing how much they know about American politics in an attempt to feel less excluded and prove themselves as more 'American' " (27). As evidence that the manipulation worked, they report that the joint effect of being Asian American and treated, relative to being white and in the control group, was that subjects listed names of an estimated 5.78 more US politicians and took approximately 88 more seconds to complete the survey (both results, $p < 0.05$).

I extend KMM's manipulation check by treating time-to-completion as a behavioral indicator across all writing tasks. Under the assumption that the list-writing prompts are the most time-intensive aspects of the survey, variation in completion time should reflect the affective and cognitive demands of those tasks. Figure A4.11 plots the relationship between total characters written and time spent on the survey by race and treatment condition. The smoothed loess curves suggest that for Asian

Americans, particularly in the treatment condition, writing output may plateau or decline at longer durations—a pattern consistent with cognitive load from the social exclusion manipulation. For white subjects, more time spent is almost linearly associated with more characters written regardless of treatment condition.

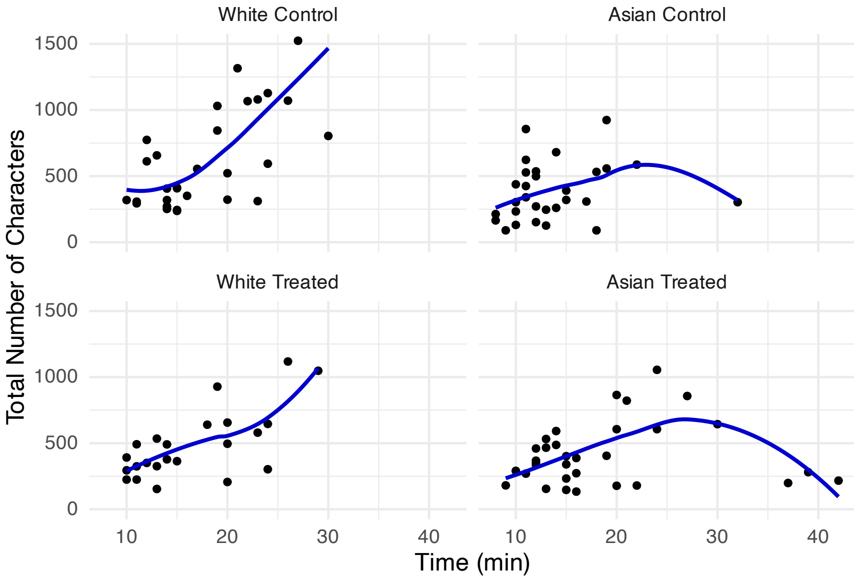

Figure A4.11. Relationship between total characters written and total survey time (minutes) by race and treatment condition. Smoothed loess curves highlight that for Asian American respondents, writing output plateaus or declines at longer durations, consistent with a potential 'writer's block' effect. Note: one white control subject who wrote more than 2000 characters is cropped for better visualization (loess curves remain unchanged).

To test for differential effects of the treatment by race on completion time, I use two methods: a Wilcoxon rank-sum test and a log-linked Gamma regression that interacts race and treatment condition. Table A4.27 presents the Wilcoxon results. The Wilcoxon test is appropriate given the relatively small sample and right-skewed completion times. As race is not randomly assigned, I use within-race comparisons to estimate the causal effect of the manipulation (Sen and Wasow 2016).

The results in Table A4.27 show that for white subjects, the social exclusion treatment does not induce a statistically significant difference in time to complete the survey ($p > 0.10$). In contrast, for Asian American subjects, the treatment does cause a statistically significant difference in time to completion ($p < 0.01$).

Group	Wilcoxon W Statistic	p-value
White Treated vs White Control	416.0	0.2045
Asian-Am Treated vs Asian-Am Control	255.5	0.0026

Table A4.27: Wilcoxon Rank Sum Test of Total Time in Study, by Race and Condition

As completion times were continuous, positive, and right-skewed, I modeled them using a log-linked Gamma regression. Table A4.28 presents the results of that analysis and shows they are consistent with the Wilcoxon rank-sum results. In a log-linked Gamma regression predicting completion time, the baseline (White respondents in the control condition) had an estimated mean of about 18 minutes. Asian American respondents completed the task roughly 25% faster than White respondents in the control condition. Among White respondents, the treatment condition was associated with a modest, non-significant 9% reduction in completion time. Among Asian American respondents who received the social exclusion treatment however, the positive and significant interaction indicates that respondents took about 50% longer than would be expected from the additive main effects alone.

	Dependent variable:
	Time (min)
Race: Asian (vs White)	–0.283*
	(0.098)
Social Exclusion Treatment	–0.095
	(0.104)
Asian × Treated	0.406*
	(0.142)
Constant	2.900*
	(0.069)
Observations	114
Log Likelihood	-356.956
Akaike Inf. Crit.	721.912
Note:	$^{*}p < 0.05$

Table A4.28: Log-linked Gamma model of time vs race and treatment condition.

These results demonstrate that metadata from writing tasks—specifically time-to-completion—can capture affective and cognitive load induced by subtle social exclusion. Even when the manipulation's effects on explicit attitudes are modest, behavioral traces from open-ended prompts can reveal underlying psychological processes.

A5. Methods: Additional Data Processing and Modeling Considerations

Working with open-ended text responses requires grappling with a range of thorny pre-processing and modeling decisions. This section outlines a set of general recommendations derived from this project but applicable to a range of related work.

1. *Data generating process*: given the wide heterogeneity in how open-ended responses are elicited and recorded, I recommend beginning any analysis of text as behavior by paying close attention to the data generating process. Survey mode (e.g., face-to-face, video, telephone, web), wave (e.g., 2016 vs. 2020 vs. 2024), and even question-level constraints (e.g., character limits) can meaningfully influence expression. Across the ANES studies, for example, web respondents in 2016 with the for/against candidate prompts were capped at 60 characters while face-to-face responses could reach over 1000. In the same survey, the maximum observed length in web responses to the four 'Most Important Problem' prompts ranged from 440 to 816 characters. In later waves, the mix of survey modes and question maxima diversified further. These structural differences can introduce floor or ceiling effects, truncation artifacts, or other forms of bias that, if not carefully assessed, may confound results.
2. *Summarize and visualize*: To mitigate risks related to idiosyncracies in the data generating process, a good starting point is to calculate and visualize basic descriptive statistics for each open-ended prompt by survey mode and wave (see, for example, Sections A1.3 and A2.1). Summary tables should be paired with visualizations (e.g., histograms, ridge plots) to reveal distributional differences and any irregularities like mass points, zero inflation, or response length cutoffs (see Section A5.1 and Figure A2.8).

3. *Transformation*: These preliminary steps also help guide the choice of data transformations. Because character counts are typically right-skewed and zero-inflated, some kind of normalization is generally necessary. For the ANES analyses that comprise most of this paper, I use the inverse hyperbolic sine (IHS) transformation. Other defensible alternatives exist, including log transforms, quantile normalization, dichotomization and discretization into ordinal bins (e.g., zero, low, medium, high).
4. *Pool*: When open-ended prompts are thematically related, such as the ANES candidate for/against questions, it may be appropriate to add and/or subtract groups of prompts by valence. In contrast, prompts like "What is the most important problem facing the country?" lack directional valence and can often be simply added. In some cases, it may be reasonable to pool raw responses first and then transform; in others, transformations may be more appropriate within mode, followed by pooling. The choice here should be guided by both empirical diagnostics and theoretical considerations, such as whether the researcher wishes to accord more weight to total verbosity (pool first, then deskew and/or normalize) or responsiveness across multiple questions (deskew and normalize first, then pool).

The difference in the order of operation is subtle but the math is straightforward: pooling raw character counts first and then deskewing with a nonlinear function will typically result in smaller values than deskewing by question and then pooling. This asymmetry is due to diminishing marginal returns of the function. Put simply, in log or IHS space, the increase from 1 to 2 is much larger than from 100 to 101. For example, imagine a hypothetical set of four open-ended prompts and a subject who responds with one five-letter word to each prompt. If we pool the character counts of all four responses and then take the log we get a value of about 3 (i.e., $\log(20) \approx 2.99$). In contrast if we take the log of each question-level character count and then pool we get a value of about 6.4 (i.e., $log(5) \times 4 = 6.44$). As shown in the analysis of Expressive Engagement in 2020 as a predictor of turnout in 2020 (see Section A2.4), there may be cases where question-level responsiveness is not the best signal.

5. *Missingness*: Decisions about missingness require careful theoretical and empirical attention. Some forms of missing data are clearly structural—like attrition from a panel study—but others fall into a gray zone between nonresponse and expressive choice. A blank response to an open-ended item might reflect refusal, avoidance, uncertainty, or disengagement. "Don't know" responses and "No" answers to filter questions may be semantically different but are often behaviorally similar to nonresponse by resulting in zero-character entries to follow-up questions. I recommend distinguishing between hard missingness (e.g., item not reached or skipped for technical reasons) and soft missingness (e.g., blank or "Don't know" responses). In some contexts, it may be substantively meaningful to treat various types of soft missingness as forms of informative nonresponse. Regardless of how these are coded, researchers should document their decisions and, ideally, show robustness to alternative coding schemes.
6. *Conceptualize*: Because any scale derived from text as behavior methods will typically be novel, I recommend using techniques like simple correlation matrices or principal component analysis (PCA) to examine dimensionality and coherence (see Sections A1.6 and A2.2).
7. *Robustness*: When incorporating new text as behavior scales into models, it is critical to assess whether they add explanatory value beyond existing closed-ended measures. For example, does Expressive Engagement improved prediction of turnout above and beyond self-reported likelihood to vote? Including such tests strengthens both the case for behavioral metadata and the clarity of the contribution.

Finally, where available, researchers might also include response time as an additional behavioral measure (see Section A4). Like character count, time-to-complete can reflect cognitive load, uncertainty, or motivation and may offer complementary insights.

A5.1 Methods: Tests for differences in Expressive Alignment and Expressive Engagement, by mode in 2016

Measure	t	df	p	mean Δ (FtF-Web)	d ≈	Interpretation
Expressive Alignment	0.92	2605	0.360	≈ 0.016	0.03	no meaningful difference
Expressive Engagement	2.99	2015	0.003	≈ 0.093	0.11	tiny, statistically detectable

Table A5.29: Mode differences (FtF – Web) for transformed metadata measures. Cohen's *d* calculated from the observed means and pooled *SD*.

Figure A5.12. Distribution of Expressive Alignment 2016 by mode.

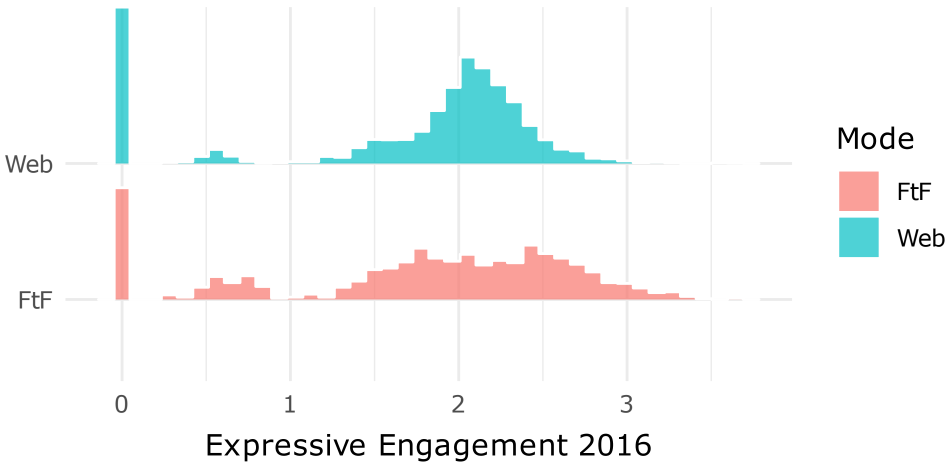

Figure A5.13. Distribution of Expressive Engagement 2016 by mode.