

**Tell Me What You Think**  
**Leveraging Open-Ended Measures in Political Psychology**

A Dissertation presented

by

**Patrick Willi Kraft**

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

**Doctor of Philosophy**

in

**Political Science**

Stony Brook University

**May 2018**

ProQuest Number:10823341

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10823341

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

Copyright by  
Patrick Willi Kraft  
2018

**Stony Brook University**  
The Graduate School

**Patrick Willi Kraft**

We, the dissertation committee for the above candidate for the  
Doctor of Philosophy degree, hereby recommend  
acceptance of this dissertation.

---

**Jennifer Jerit – Dissertation Advisor**  
Professor, Department of Political Science

---

**Yanna Krupnikov – Chairperson of Defense**  
Associate Professor, Department of Political Science

---

**Stanley Feldman – Committee Member**  
Professor, Department of Political Science

---

**Michael Peress – Committee Member**  
Associate Professor, Department of Political Science

---

**Arthur Spirling – Outside Committee Member**  
Associate Professor of Politics and Data Science  
Department of Political Science, New York University

This dissertation is accepted by the Graduate School

---

**Charles Taber**  
Dean of the Graduate School

Abstract of the Dissertation

**Tell Me What You Think  
Leveraging Open-Ended Measures in Political Psychology**

by

**Patrick Willi Kraft**

**Doctor of Philosophy**

in

**Political Science**

Stony Brook University

**2018**

Though verbally expressing attitudes is one of the most ubiquitous ways people engage in politics, this basic feature of political life is rarely studied directly. Building on recent advances in automated text analysis, I develop new measures to systematically examine verbatim political attitude expression. By analyzing how citizens describe their beliefs and discuss them with peers, my research advances previous theoretical insights on the nature of political sophistication as well as the role of morality in politics and persuasion. The first part of the dissertation shows that the complexity with which people discuss political preferences, or their discursive sophistication, is a better predictor of political competence than factual knowledge alone. My measure of discursive sophistication furthermore suggests that—in contrast to previous findings in the literature—women are by no means less politically sophisticated than men. In the second part, I examine ideological differences in the contents of expressed attitudes. The analyses reveal systematic variation in the use of moral language between liberals and conservatives when talking about politics, a finding that is consistent with previous research in moral psychology. However, the reliance on morality is influenced by the degree to which people are exposed to moral rhetoric in the media. The third part of the dissertation investigates how the expression of moral considerations affects persuasion and attitude change in the context of online discussions. While moral appeals do not change people’s minds across the board, those who hear arguments that are morally congruent with their preexisting attitudes are more likely to be persuaded. Overall, the dissertation advocates for a greater use of text-as-data and open-ended measures in the area of political psychology.

*For my brother,  
who began this journey with me  
but was unable to see me complete it.*

# Contents

<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xii</b>
<b>1 The Politics of Attitude Expression</b>	<b>1</b>
<b>2 Let's Talk Politics</b>	<b>6</b>
2.1 Introduction . . . . .	7
2.2 Opinion Formation and Attitude Expression . . . . .	9
2.3 Measuring Discursive Sophistication . . . . .	10
2.4 An Overview of Data Sources and Open Ended Items . . . . .	13
2.4.1 2012 & 2016 American National Election Study . . . . .	13
2.4.2 2015 YouGov Survey . . . . .	14
2.4.3 Swiss Referendum Survey . . . . .	15
2.5 A First Look at Discursive Sophistication . . . . .	15
2.6 Discursive Sophistication and Political Competence . . . . .	18
2.6.1 Engagement and Participation in Politics . . . . .	18
2.6.2 Incorporation of New Information . . . . .	20
2.6.3 Precise Positioning of Candidates . . . . .	21
2.6.4 Well-Justified Policy Preferences . . . . .	22
2.7 Application: The Gender Gap in Political Knowledge . . . . .	25
2.7.1 Descriptive Results . . . . .	26
2.7.2 Controlling for Alternative Explanations . . . . .	27
2.7.3 Explaining the (Lack of the) Gender Gap . . . . .	29
2.8 Conclusion . . . . .	30

<b>3</b>	<b>Is It All About Values?</b>	<b>32</b>
3.1	Introduction . . . . .	33
3.2	Method . . . . .	34
3.3	Results . . . . .	36
3.3.1	Ideological Differences . . . . .	36
3.3.2	Moral Considerations and Vote Choice . . . . .	38
3.3.3	Media Content and Exposure to Moral Rhetoric . . . . .	39
3.3.4	Robustness Checks . . . . .	40
3.4	Discussion . . . . .	41
<b>4</b>	<b>Change My View</b>	<b>43</b>
4.1	Introduction . . . . .	44
4.2	Theoretical Background . . . . .	46
4.2.1	Two Routes to Persuasion . . . . .	46
4.2.2	Morality and the Potential for Compromise . . . . .	48
4.2.3	Hypotheses . . . . .	50
4.3	The Subreddit “ChangeMyView” . . . . .	51
4.4	Opinion Change in Online Discussions . . . . .	54
4.5	What Makes an Argument Persuasive? . . . . .	59
4.5.1	Moral Appeals are Futile... . . . .	62
4.5.2	...Unless We’re Speaking the Same Moral Language . . . . .	63
4.6	Conclusion . . . . .	65
<b>5</b>	<b>Concluding Remarks</b>	<b>68</b>
<b>6</b>	<b>Bibliography</b>	<b>70</b>
<b>A</b>	<b>Supplementary Material for Chapter 2</b>	<b>79</b>
A.1	Information on Discursive Sophistication Components . . . . .	80
A.1.1	Distribution of Word Counts in Open-Ended Responses . . . . .	80
A.1.2	Overview of Topic Proportions . . . . .	81
A.1.3	Discursive Sophistication Components . . . . .	84
A.2	Pre-Processing and Topic Model Specification . . . . .	85
A.2.1	PreText Analysis . . . . .	85
A.2.2	Robustness Checks for Varying Model Specifications . . . . .	86



A.3	Additional Information on Remaining Variables . . . . .	88
A.3.1	Item Selection and Recoding . . . . .	88
A.3.2	Variable Distributions – 2012 ANES . . . . .	90
A.3.3	Variable Distributions – 2016 ANES . . . . .	92
A.3.4	Variable Distributions – 2015 YouGov . . . . .	94
A.4	Robustness Checks . . . . .	95
A.4.1	Controlling for Personality Characteristics . . . . .	95
A.4.2	Controlling for Individual Response Length . . . . .	96
A.5	Tables of Model Estimates . . . . .	97
A.5.1	Main Analyses . . . . .	97
A.5.2	Robustness Checks . . . . .	104
<b>B</b>	<b>Supplementary Material for Chapter 3</b>	<b>108</b>
B.1	Moral Foundations Dictionary . . . . .	109
B.2	Information on Data, Variables, and Recoding . . . . .	111
B.2.1	Open-ended Responses and MFT Scores in 2012 ANES . . . . .	111
B.2.2	Moralization in Individual Media Environments . . . . .	114
B.2.3	Remaining Variables in 2012 ANES . . . . .	116
B.2.4	Open-ended Responses and MFT Scores in Replication Sample . . . . .	118
B.2.5	Control Variables in Replication Sample . . . . .	119
B.3	Additional Model Results & Robustness Checks . . . . .	120
B.3.1	Replicating Ideological Differences using RDD Adult Sample . . . . .	120
B.3.2	Negations and Valence in Open-Ended Responses (2012 ANES) . . . . .	122
B.3.3	MFT and Party/Candidate Evaluations (2012 ANES) . . . . .	124
B.3.4	Other Correlates of General Moralization in Attitude Expression . . . . .	125
B.3.5	Comparing General Media MFT Scores with Manual Coding . . . . .	126
B.3.6	Face Validity of Sample Responses . . . . .	128
B.4	Tables of Model Estimates . . . . .	129
B.4.1	Ideological Differences . . . . .	129
B.4.2	Moral Considerations and Vote Choice . . . . .	130
B.4.3	Media Content and Exposure to Moral Rhetoric . . . . .	131
B.4.4	Additional Model Results & Robustness Checks . . . . .	132
<b>C</b>	<b>Supplementary Material for Chapter 4</b>	<b>142</b>
C.1	Basic Rules on the Subreddit ChangeMyView . . . . .	143

C.2	Moral Foundations Dictionary . . . . .	144
C.3	Distribution of Moral Foundation Proportions in Paired Data . . . . .	146
C.4	Structural Topic Model Results . . . . .	147
C.4.1	Original Posts . . . . .	147
C.4.2	Responses Challenging the OP . . . . .	149
C.5	Tables of Model Estimates . . . . .	151

## List of Figures

2.1	Correlation matrix of conventional political knowledge metrics and discursive sophistication . . . . .	16
2.2	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2012 and 2016 ANES . . . . .	19
2.3	Expected information retrieval in the 2015 YouGov Study as a function of political sophistication . . . . .	20
2.4	Error variance reduction in candidate placements on multiple issues in the 2012 and 2016 ANES . . . . .	23
2.5	Discursive sophistication and manually coded level of justification in Swiss post-referendum surveys . . . . .	24
2.6	The gender gap in political sophistication . . . . .	27
2.7	Common determinants of political sophistication . . . . .	28
2.8	Gender differences in topic proportions in open-ended responses . . . . .	30
3.1	Difference between liberals and conservatives in moral considerations in open-ended responses . . . . .	37
3.2	Effect of moral considerations on predicted probabilities of voting for the Democratic rather than Republican candidate conditional . . . . .	38
3.3	Effect of moral content in individual media environments on moral reasoning in open-ended responses . . . . .	40
4.1	Average topic proportions in opening statements on /r/ChangeMyView/ . . . . .	56
4.2	Number of original posts on /r/ChangeMyView/ that resulted in opinion change . . . . .	57
4.3	Moral foundations and persuadability . . . . .	58
4.4	Number of opening statements in the paired argument data that included <i>any</i> term mentioned in the MFT dictionary . . . . .	60
4.5	Difference in response lengths between successful and unsuccessful counterarguments . . . . .	61

4.6	Moral foundations and persuasiveness . . . . .	62
4.7	Moral congruence and persuasiveness . . . . .	64
A.1	Histograms of total word count in open-ended responses . . . . .	80
A.2	Estimated topic proportions in the 2012 and 2016 ANES based on the structural topic model . . . . .	81
A.3	Estimated topic proportions in the 2015 YouGov survey based on the structural topic model . . . . .	82
A.4	Estimated topic proportions in the Swiss survey based on the structural topic model	83
A.5	Correlation matrix of individual components of discursive sophistication. . . . .	84
A.6	PreText analysis of pre-processing decisions of open-ended responses across all datasets . . . . .	85
A.7	Robustness of discursive sophistication measure for different pre-processing choices and topic model specifications . . . . .	87
A.8	Histograms of dependent variables included in 2012 ANES . . . . .	90
A.9	Histograms of independent variables included in 2012 ANES . . . . .	91
A.10	Histograms of dependent variables included in 2016 ANES . . . . .	92
A.11	Histograms of independent variables included in 2016 ANES . . . . .	93
A.12	Histogram of variables included in 2015 YouGov survey . . . . .	94
A.13	Effects of sophistication on political engagement controlling for personality characteristics in the 2012 and 2016 ANES . . . . .	95
A.14	Effects of sophistication on political engagement controlling for response length in the 2012 and 2016 ANES . . . . .	96
B.1	Histograms displaying the distribution of individual response lengths in number of words for each respective item category . . . . .	111
B.2	Proportion of respondents mentioning each of the moral foundations in any of their open-ended responses (2012 ANES) . . . . .	112
B.3	Weights for individual MFT dictionary terms . . . . .	113
B.4	General MFT scores for media sources during 2012 U.S. Presidential campaign .	114
B.5	Histogram of moralization in individual media environments . . . . .	115
B.6	Histograms of variables included in analyses . . . . .	117
B.7	Proportion of respondents mentioning each of the moral foundations in any of their open-ended responses (RDD adult sample) . . . . .	118
B.8	Histograms of variables included in replication survey . . . . .	119
B.9	Replication of main model predicting MFT scores based on ideology using RDD adult sample . . . . .	120

B.10	Replication of main model predicting MFT scores based on ideology by different subgroups . . . . .	122
B.11	Effect of MFT score on predicted feeling thermometer differential . . . . .	124
B.12	Effect of moral media content on moral reasoning in open-ended responses controlling for political interest measures . . . . .	125
B.13	Validity check comparing general MFT Scores with manual coding of newspaper articles . . . . .	126
B.14	Validity check comparing individual MFT Scores with manual coding of newspaper articles . . . . .	127
C.1	Moral Foundations in Paired Data . . . . .	146
C.2	Average topic proportions in opening statements on CMV based on structural topic model . . . . .	147
C.3	Differences in topic proportions between opening statements on CMV that resulted in opinion change versus not . . . . .	148
C.4	Average topic proportions in posts challenging the OP on CMV based on structural topic model . . . . .	149
C.5	Differences in topic proportions between persuasive and non-persuasive posts challenging the OP on CMV . . . . .	150

## List of Tables

2.1	Example of open-ended responses for low and high scores on discursive sophistication	17
A.1	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2012 ANES	97
A.2	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2016 ANES	98
A.3	Effects of sophistication on information retrieval in the 2015 YouGov study	99
A.4	Error variance reduction in candidate placements on multiple issues in the 2012 ANES	100
A.5	Error variance reduction in candidate placements on multiple issues in the 2016 ANES	101
A.6	Effects of gender on discursive sophistication and factual knowledge in the 2012 ANES and 2016 ANES	102
A.7	Effects of gender on discursive sophistication and factual knowledge in the 2015 YouGov Study	103
A.8	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for personality characteristics in the 2012 ANES	104
A.9	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for personality characteristics in the 2016 ANES	105
A.10	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for individual response length in the 2012 ANES	106
A.11	Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for individual response length in the 2016 ANES	107

B.1	Missing open-ended responses . . . . .	111
B.2	Sample of open-ended responses in the 2012 ANES . . . . .	128
B.3	Tobit models predicting MFT score for each foundation based on ideology . . . . .	129
B.4	Logit models predicting democratic vote choice based on MFT score for each foundation . . . . .	130
B.5	Tobit model predicting overall reliance on moral foundations (sum of MFT scores) based on media moralization, political knowledge, media exposure, and frequency of political discussions . . . . .	131
B.6	Tobit models predicting MFT score for each foundation based on ideology (telephone survey replication) . . . . .	132
B.7	Virtues only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	133
B.8	Vices only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	134
B.9	Likes only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	135
B.10	Dislikes only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	136
B.11	Democratic party/candidate only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	137
B.12	Republican Party/Candidate only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	138
B.13	In-party only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	139
B.14	Out-party only: Tobit models predicting MFT score for each foundation based on ideology . . . . .	140
B.15	OLS models predicting feeling thermometer differentials based on MFT score for each foundation . . . . .	141
C.1	Logit models predicting argument persuasiveness as a function of moral word use	151
C.2	Logit models predicting argument persuasiveness as a function of moral congruence with OPs' opening statements . . . . .	152

## Acknowledgements

Throughout graduate school and the process of writing this dissertation, I was incredibly lucky to be surrounded by inspiring mentors and great friends who made the completion of this project possible. I want to thank my advisor, Jennifer Jerit, who taught me how to turn (sometimes vague) research ideas into polished articles. She provided steadfast support, invaluable feedback, and always reminded me to think about the bigger puzzle. I also am deeply grateful to the other members of my dissertation committee, Stanley Feldman, Yanna Krupnikov, Michael Peress, and Arthur Spirling. They all challenged me to think more critically about the psychology of politics as well as research methodology and statistics. They helped me grow intellectually through many insightful conversations during and after the courses they taught me. To name just one example, Stanley Feldman's incredible class on political ideology became the starting point for this dissertation.

Many others helped me along the way. I originally came to Stony Brook, for what I believed to be an exchange year, to work with Milton Lodge and Chuck Taber—and both quickly convinced me to stay for the PhD. Matt Lebo provided incomparable support from my first day at the department and taught me how to efficiently handle a steadily growing number of projects. I also benefited greatly from valuable discussions with Hannah Nam (and the entire Nam lab!) as well as Rüdiger Schmitt-Beck, Thorsten Faas, and Thomas Gschwend, who guided me through my initial years at the University of Mannheim and thereby paved my way to Stony Brook.

I am especially thankful for all the amazing and inspiring graduate students that I had the privilege to learn from—both in Stony Brook and Mannheim—many of whom became close friends. While I can't possibly list everyone here, I am indebted to Fridolin Linder for his friendship, as well as Alexa Bankert for her support through the first (at times difficult) years of graduate school. Furthermore, I want to thank David Stack for exceptionally productive research meetings in the graduate student lounge.

Beyond the department of political science, I am grateful to Adrián Soto Cambres, Tristan Delaney, Ludwig Krinner, Hans Niederhausen, and Kent Henderson. You became my family away from home. Lastly, I thank Talbot Andrews for making the last stretch of finishing the dissertation the most enjoyable period of my entire time in graduate school. Thank you so much to all of you, I could have never done it without you.



## Introduction: The Politics of Attitude Expression

### Abstract

Though verbally expressing attitudes is one of the most ubiquitous ways people engage in politics, this basic feature of political life is rarely studied directly. Building on recent advances in automated text analysis, I develop new measures to systematically examine verbatim political attitude expression. By analyzing how citizens describe their beliefs and discuss them with peers, my research advances previous theoretical insights on the nature of political sophistication as well as the role of morality in politics and persuasion. The first part of the dissertation shows that the complexity with which people discuss political preferences, or their discursive sophistication, is a better predictor of political competence than factual knowledge alone. My measure of discursive sophistication furthermore suggests that—in contrast to previous findings in the literature—women are by no means less politically sophisticated than men. In the second part, I examine ideological differences in the contents of expressed attitudes. The analyses reveal systematic variation in the use of moral language between liberals and conservatives when talking about politics, a finding that is consistent with previous research in moral psychology. However, the reliance on morality is influenced by the degree to which people are exposed to moral rhetoric in the media. The third part of the dissertation investigates how the expression of moral considerations affects persuasion and attitude change in the context of online discussions. While moral appeals do not change people’s minds across the board, those who hear arguments that are morally congruent with their preexisting attitudes are more likely to be persuaded. Overall, the dissertation advocates for a greater use of text-as-data and open-ended measures in the area of political psychology.

*“What is important to study cannot be measured  
and that what can be measured is not important to study.”*

*Converse (1964, 206)*

Politics is, at its core, a social and communicative process: government officials hold speeches to lay out their agenda, legislators address parliaments in support of their bills, and candidates engage in public debates prior to elections. Given that political processes routinely generate large amounts of text—such as manifestos, transcripts, or laws—it seems natural that there is a growing body of research in political science that is eager to incorporate new approaches from automated text analysis. For example, there have been numerous studies that analyzed party documents to make inferences about their ideological positions (Laver and Garry, 2000; Laver, Benoit, and Garry, 2003; Slapin and Proksch, 2008; Benoit, Laver, and Mikhaylov, 2009). Other researchers focused on legislative speeches and related documents to examine parliamentarians’ ideology (Diermeier et al., 2012), broader legislative agendas (Quinn et al., 2010; Grimmer, 2010), government responsiveness (Eggers and Spirling, 2014, 2018), elite polarization (Jensen et al., 2012), and democratization (Spirling, 2016). However, virtually all studies in this area exclusively focus on elite communications or text produced by political institutions.

Of course, political communication does not only play an integral role on the elite level. Citizens regularly discuss political issues with their co-workers, neighbors, friends, and family. In fact, political conversations are among the main sources of information about politics for individuals (Huckfeldt et al., 1995; Beck et al., 2002; McClurg, 2006). Furthermore, a multitude of studies have demonstrated how attitudes and behavior are influenced through interactions in social networks (Mutz, 2002; McClurg, 2003; Lazer et al., 2010; Ahn, Huckfeldt, and Ryan, 2010). Whether it occurs in personal conversations, on social media, or when citizens contact government officials—people routinely express their political ideas, attitudes, and preferences in their own words. In this dissertation, I explore new ways to leverage these verbatim attitude expressions for the study of politics—by directly analyzing how people describe their preferences in open-ended survey questions or how they try to persuade each other in the context of discussions.

Although verbatim attitude expression is ubiquitous as a medium for conveying political views, most public opinion research is surprisingly hesitant to directly examine how individuals articulate their beliefs. Instead of exploring open-ended survey responses to study political preference formation among the public, scholars largely rely on conventional closed items. While there are some notable exceptions involving qualitative analyses of individual attitudes (e.g. Chong, 1993), quantitative researchers in political psychology and related fields lacked the necessary tools to fully leverage large amounts of text data—for instance in the context of survey research—without involving manual coders. To be fair, the discipline’s reluctance to incorporate text-based approaches in the study of political attitudes is not only due to the absence of appropriate tools and methods. After all, there are good theoretical reasons for scholars to rely on their established survey instruments rather than analyzing open-ended responses. Common concerns include that respondents might be more reluctant to provide answers to open-ended items, or response patterns may be confounded by people’s ability to articulate a coherent response. However, these doubts appear less warranted than initially thought (e.g., Krosnick, 1999).

In theory, a survey can be characterized as a formalized conversation between two individuals that is not unlike other social encounters (c.f. Sudman, Bradburn, and Schwarz, 1996, see also Grice 1975, 1978). The most important difference between open-ended and closed survey questions in this context is the flexibility with which an interviewee can formulate his or her response to a given question. The open-ended setting is ultimately less formalized and can therefore resemble other types of conversations more closely. While this flexibility creates significant challenges for the researcher, it also provides great potential to make inferences about political attitudes that could not be captured in closed survey questions. For example, open-ended items can help us understand which considerations related to an attitude object is salient for an individual, without explicitly priming a specific evaluative category (e.g. Geer, 1988, 1991). Furthermore, if open-ended responses reveal the considerations currently salient and accessible for an individual, they can provide insights about framing effects and related phenomena (Nelson, Oxley, and Clawson, 1997; Chong and Druckman, 2007). More generally, the considerations people raise in a flexi-

ble context of open-ended allows us to make inferences about people's ambivalence in political attitudes (Basinger and Lavine, 2005; Lavine, Johnston, and Steenbergen, 2012).

A skeptic might argue that the content of individual utterances about politics are nothing but rationalizations of underlying affective predispositions (e.g., Lodge and Taber, 2013). Nevertheless, language remains the main medium for people to convey their attitudes, beliefs, and preferences to others. Explicit attitude expression through language is the foundation of any communication process and therefore a crucial mediator of social influence. As such, open-ended responses are informative insofar as they represent the way individuals describe and justify their political attitudes. Granted, there may be advantages in using closed measures to answer certain research questions. However, this does not imply that open-ended measures should be ignored altogether.

Overall, the goal of this dissertation is to show how the incorporation of open-ended measures provides new and promising avenues to test important hypotheses related to political cognition, attitude formation, and social influence. The dissertation thereby not only aims to contribute to the literature in political psychology, but also to the field of quantitative text analysis. In taking a closer look at the underlying psychological and cognitive mechanisms that guide response behavior and communication, this dissertation can inform new approaches in text analysis, which have been largely devoid of a cognitive theory guiding the production of political text. The first empirical chapter demonstrates how open-ended responses about political preferences can be utilized to measure political sophistication. The proposed measure—discursive sophistication—proves to be a better predictor of competence in the realm of politics than conventional fact-based measures of political knowledge. The second empirical chapter examines ideological differences in moral reasoning and demonstrates how open-ended responses can be leveraged to show how media exposure can increase the moralization of politics. Lastly, the third empirical chapter analyzes the content of online discussions to investigate the role of moral appeals in persuasion and compromise. In contrast to some previous research in moral psychology, the results indicate that moralized arguments do not necessarily foster divisiveness. Instead, they may help overcome

disagreements as long as people speak the same moral language.

Converse (1964) began his seminal study on the nature of belief systems by postulating that “what is important to study cannot be measured.” Yet, recent advances in automated text analysis may provide new avenues to capture concepts that are indeed important to study and have been difficult to assess in the past. This dissertation attempts to highlight such areas where the incorporation of open-ended measures can contribute to long-standing debates in the literature. From a methodological perspective, the dissertation discusses the advantages—and potential pitfalls—of open ended measures in the field of political psychology. If we aim for a better understanding of how people *think* about politics, we should focus more on how they *talk* about politics.

## Chapter 2

### Let's Talk Politics:

### A Naive Approach for Measuring Political Sophistication\*

#### Abstract

This paper proposes a simple but powerful framework to measure political sophistication based on open-ended survey responses. *Discursive sophistication* utilizes automated text analysis methods to capture the complexity of individual attitude expression. I validate the approach by comparing it to conventional political knowledge metrics in multiple studies using different batteries of open-ended items. The paper then illustrates how the measure can help refine previous insights from the literature such as the oft-cited gender gap in political knowledge. Women might know fewer facts about institutions and elites, but they do not differ substantively in the sophistication of their expressed political beliefs.

---

\*The code for this chapter is available on GitHub: <https://github.com/pwkraft/knowledge>.

### 2.1 Introduction

One of the most important tasks for citizens in modern democracies is to vote for candidates who represent their interests and hold their elected officials accountable. While there have been long-standing debates about whether citizens are sufficiently informed to fulfill this task, fundamental issues regarding the measurement of knowledge continue to plague the discipline (Mondak, 2001; Sturgis, Allum, and Smith, 2008; Pietryka and MacIntosh, 2013). Most analyses rely on batteries that assess individuals' factual knowledge about political institutions and officeholders (e.g., Delli Carpini and Keeter, 1996). Theoretically, these survey questions should cover information that is necessary and/or sufficient for citizens to make competent decisions in a given context (c.f., Lupia, 2006, 2015). Yet, determining such a set of items proves to be extremely difficult, especially since there are systematic differences in types of knowledge (Barabas et al., 2014). Even within a given policy area, people may disagree about which facts are crucial for political competence due to inherent value differences (Lupia, 2015).

Despite these difficulties, most empirical studies rely on a set of off-the-shelf knowledge questions rather than justifying their choices theoretically. As Lupia (2006, 219) points out, “[m]ost political knowledge questions are not derived from a replicable or transparent logic about how their answers bear on a voter’s ability to make decisions of a particular quality.” It is therefore not surprising that conventional metrics do not properly capture policy-specific information (e.g., Gilens, 2001) or other knowledge relevant to citizens’ preferences and decision-making (c.f., Graber, 2001, 43–68). In a recent review, Cramer and Toff (2017, 756) eloquently summarize: “All of this work suggests that we are missing a lot by equating information levels as measured in traditional knowledge batteries with civic competence. By focusing on what people do not know rather than what they do know and how they use that information, we are likely missing the empirical reality of citizens’ political knowledge.”

While some scholars contend that simple tests of factual information are nevertheless the best available proxy for political awareness (e.g., Zaller, 1990), others rely on broader conceptualiza-

tions of sophistication that incorporate additional dimensions such as education and income (e.g., Jacoby, 2006). Notwithstanding, the fundamental issue remains that knowledge quizzes rarely cover information that is relevant for citizen competence (Lupia, 2006). In a similar vein, Druckman (2014) describes individual levels of political information as inadequate to measure “quality opinion” since there is no consensus about what information is necessary in the first place. Instead, Druckman advocates “less focus on the *content/substance* of opinions [...] and *more* on the *process* and specifically the *motivation* that underlies the formation of those opinions” (2014, 478, emphasis in the original).

The framework proposed herein follows this call by using a person’s open-ended responses to develop an indicator of *discursive sophistication*. The measure examines how respondents discuss their political beliefs in their own words and incorporates information about the number of considerations raised, the relative descriptiveness in word choice, as well as the level of opinionation. The approach is therefore *naive* in that it does not presuppose pieces of information as necessary for political competence but rather examines the respondents’ justification of their preferences at face value. Measuring sophistication based on people’s verbatim attitude expression provides two major advantages compared to off-the-shelf factual knowledge items: (1) it directly captures the extent to which a respondent’s political beliefs are based on elaborate reasoning, and (2) it can easily pinpoint competence in specific areas by incorporating targeted open-ended items.

I validate the measure across multiple data sets by comparing it to conventional factual knowledge scores as predictors of various indicators of competence. While the measures share a considerable amount of variance, they are far from equivalent. Indeed, discursive sophistication is a stronger predictor of turnout and other forms of political participation than traditional metrics. After validating the measurement approach, the paper illustrates how discursive sophistication can help refine previous insights in the literature by re-examining an oft-cited finding in empirical research—the gender gap in political knowledge. Contrary to previous research, I find no evidence for such a gap based on open-ended responses. While women might score lower than men on factual knowledge about political institutions and elites, there are no differences in the complexity



of expressed political attitudes.

### 2.2 Opinion Formation and Attitude Expression

In modern democracies, citizens can engage in politics through various means such as voting in local, state, or federal elections. Depending on the institutional setup, they may also directly decide on specific policies through referenda. In these contexts, we are concerned with the ability of citizens to make high quality decisions in accordance with their underlying interests. Given that it is challenging to determine what information is indeed necessary and/or sufficient to make competent decisions in a given context, a useful alternative is to concentrate on whether people are motivated to engage in elaborate reasoning when forming their preferences (Druckman, 2014). In previous research, scholars have induced people to engage in in-depth processing by asking them to *justify* their opinions (e.g., by providing specific reasons; Kunda and Sinclair, 1999; Redlawsk, 2002; Bolsen, Druckman, and Cook, 2014a; Druckman, 2014). In an analogous way, we can examine *how* citizens justify their preferences in order to evaluate whether they engaged in elaborate and sophisticated reasoning (see also Rosenberg, 1988; Rosenberg, Ward, and Chilton, 1988). If respondents are motivated and able to engage in in-depth processing to form quality opinions, they should discuss multiple considerations related to a political issue and show awareness of arguments for and against certain positions (Cappella, Price, and Nir, 2002). Rather than trying to develop recall items that presupposes a set of facts as necessary for political competence, I therefore analyze *how* individuals discuss their preferences related to a given political task.<sup>1</sup>

My approach is consistent with influential theoretical accounts of political sophistication which focus on the *structure* of belief systems. For example, Converse (1964) emphasizes the importance of the level of conceptualization as the main characteristic of sophistication rather than isolated

---

<sup>1</sup>A similar approach is taken by Colombo (2016) who investigates the competence of Swiss citizens voting in policy referenda. Colombo conceptualizes competence as a voter's ability to justify his or her political decisions, and measures the concept by manually coding open-ended responses to survey questions.

pieces of factual information. Similarly, Tetlock (1983, 1993) uses the term *integrative complexity* to describe the degree to which considerations related to an issue are interconnected. Luskin (1987) also defines political sophistication based on the structure of individual belief systems, arguing that they can vary on three separate dimensions: (1) their *size* – i.e. the number of cognitions, (2) their *range* – i.e. the dispersion of cognition over categories, and (3) their *constraint* – i.e. the extent to which cognitions are interconnected in a meaningful way. Political sophistication, in turn, is seen as the conjunction of these dimensions: “A person is politically sophisticated to the extent to which his or her [political belief system] is large, wide-ranging, and highly constrained.” (Luskin, 1987, 860).

Overall, this body of work suggests that differences in sophistication should be reflected in the way individuals describe and justify their political beliefs. Crucially, a measure of sophistication that is based on how individuals discuss their preferences in their own words can be directly applied in various settings to target specific political tasks such as choosing between candidates, parties, or policy propositions. Rather than having to devise a new set of questions that attempt to capture information necessary to make competent decisions, we can simply analyze how respondents elaborate on their related preferences in verbatim.

## 2.3 Measuring Discursive Sophistication

How would a politically sophisticated person who engages in in-depth processing discuss his or her views compared to a less informed individual? Consider a survey where respondents are asked to describe their attitudes toward specific policies or candidates running for office in a set of open-ended items. In such a scenario, the structure of individual political belief systems (i.e., size, range, and constraint) as well as the level of motivation to engage in more elaborate reasoning should be reflected in their verbatim responses. In the following, I discuss three different attributes of open-ended survey responses that should be indicative of sophistication in attitude expression.

First of all, sophisticated individuals should be able to elaborate more on their political at-

titudes. If people possess a large, wide-ranging, and constrained belief system, they should be able to recall a large number of *considerations* related to political actors or issues. I rely on the structural topic model framework (Roberts et al., 2014) to extract the number of topics mentioned by each respondent in a survey.<sup>2</sup> First, denote  $\mathcal{W}_i$  as the set of words contained in a response of individual  $i$ . Each word  $w \in \mathcal{W}_i$  is assigned to a topic  $t^* \in \{1, \dots, T\}$ , such that  $P(t^*|w, X_i) > P(t|w, X_i) \forall t \neq t^*$ .<sup>3</sup> In other words, each unique term in a response is assigned to the topic that has the highest likelihood of having generated that term, given the model. The set of topics that are mentioned by respondent  $i$  across all words in  $\mathcal{W}_i$  can then be denoted as  $\mathcal{T}_i^*$  and the number of considerations can be written as:

$$\text{considerations}_i = \frac{|\mathcal{T}_i^*|}{\max_i |\mathcal{T}_i^*|}. \quad (2.1)$$

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

However, sophisticated respondents should not only be able to mention a larger number of raw considerations when discussing politics. The level of sophistication should also be reflected in the *word choice* describing the underlying issues. Individuals who possess a constrained system of beliefs should be more inclined to use terms that are highly descriptive of a given topic (e.g., the *economy* or *taxes*) rather than broad terms that could be attributed to any topic and are not clearly related to politics. Highly descriptive word choice is conceptualized as the sum of term likelihoods  $P(w|t^*)$  given topic assignments over the entire set of words in  $\mathcal{W}_i$ :

$$\text{word choice}_i = \frac{\sum_{\mathcal{W}_i} P(w|t^*)}{\max_i [\sum_{\mathcal{W}_i} P(w|t^*)]} \quad (2.2)$$

Again, the measure is re-scaled to range from zero to one by dividing all values by the empirical

<sup>2</sup>See below for more information on the set of open-ended responses, pre-processing choices, as well as on the topic model specification.

<sup>3</sup>Note that  $P(t|w, X_i) = \frac{P(w|t)P(t|X_i)}{P(w|X_i)}$ . In the context of structural topic models,  $X_i$  denotes the covariates used to predict individual topic prevalence (see Roberts et al., 2014, for details).

## 2.3. MEASURING DISCURSIVE SOPHISTICATION

maximum observed across all individuals in the data.

Lastly, sophisticated individuals should hold opinions about each political actor or policy that they are asked to discuss. Given a set of multiple open-ended probes focusing on different issues, sophisticates should be able to express their attitudes towards each question in terms of both approval or disapproval. Responses that reflect high levels of sophistication should therefore display a greater level of *opinionation*, which is conceptualized as the diversity of relative lengths for each open-ended response (specified as the Shannon entropy):

$$\text{opinionation}_i = \frac{-\sum_{j=1}^J p_{ij} \ln p_{ij}}{\ln J} \quad (2.3)$$

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

Together, the three measures form a composite metric of sophistication in political attitude expression by calculating their respective average for each respondent. Like each individual component, the resulting *discursive sophistication* score ranges from 0 to 1:

$$\text{discursive sophistication}_i = \frac{1}{3}(\text{considerations}_i + \text{word choice}_i + \text{opinionation}_i). \quad (2.4)$$

Overall, a highly sophisticated individual can be expected to respond to a set of open-ended items by giving a more elaborate response that focuses on multiple considerations using terms that are highly descriptive of each topic and addresses his or her attitudes towards all relevant political actors or policies more or less equally.<sup>4</sup>

<sup>4</sup>Note that this approach differs from recent work on sophistication in speeches and other sources of political texts (e.g., Spirling, 2016; Benoit, Munger, and Spirling, 2017) as it explicitly targets complexity independent of pure linguistic style.

### 2.4 An Overview of Data Sources and Open Ended Items

The measure of discursive sophistication is validated using multiple surveys employing different sets of open-ended questions. Each survey focuses on sophistication in the context of distinct political tasks, namely the evaluation of (1) candidates running for public office, (2) broad issue areas such as health care and gun legislation, and (3) policy referenda. The data sets and items used to compute discursive sophistication are briefly described below.<sup>5</sup>

#### 2.4.1 2012 & 2016 American National Election Study

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

---

<sup>5</sup>See Appendix A.1 for descriptive information on open-ended responses in each dataset, structural topic model results, and individual components of discursive sophistication. Appendix A.2 contains further details on pre-processing steps and modeling choices for the structural topic models as well as robustness checks, which include preText analyses proposed by Denny and Spirling (2018). Lastly, Appendix A.3 provides information on the remaining variables included in the analyses.

### 2.4.2 2015 YouGov Survey

In order to replicate and extend the main analyses, I rely on a separate nationally representative survey employing an alternative set of open-ended responses. The data was collected by YouGov in December 2015 and contains responses of 1000 U.S. citizens.<sup>6</sup> As part of this study, respondents were asked to describe their attitudes towards two prominent political issues that were discussed frequently in the media. First, they were asked in a closed format whether they favor or oppose stricter gun laws. Subsequently, they were asked to respond to the following two questions:

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

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

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

Here, discursive sophistication is computed based on the verbatim responses to the four preceding questions using the same procedures described above. Compared to the open-ended likes/dislikes items included in the 2012 and 2016 ANES, the questions directly address considerations related to specific policy issues that were prominent in the political discourse at the time of the survey.

<sup>6</sup>See Clifford and Jerit (2018) for details on the study.

Respondents who did not provide an answer to any of the open-ended questions were removed from the analysis (48).

### 2.4.3 Swiss Referendum Survey

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

## 2.5 A First Look at Discursive Sophistication

Before turning to the validation, I begin by directly comparing discursive sophistication to alternative metrics of political knowledge in the 2012 and 2016 ANES. The standard approach to measuring political knowledge in surveys is to ask a set of factual questions about political institutions. The ANES surveys include such a basic item battery, inquiring for example about the number of times an individual can be elected President of the United States, or how the current U.S. federal budget deficit compares to the deficit in the 1990s. I combine responses on these items to form an additive index of *factual knowledge* about politics. As an additional benchmark, I consider *interviewer assessments* of each respondent's political sophistication (c.f., Bartels 2005

## 2.5. A FIRST LOOK AT DISCURSIVE SOPHISTICATION

for an example of a study that relies on interviewer assessments; but see also Ryan 2011).<sup>7</sup>

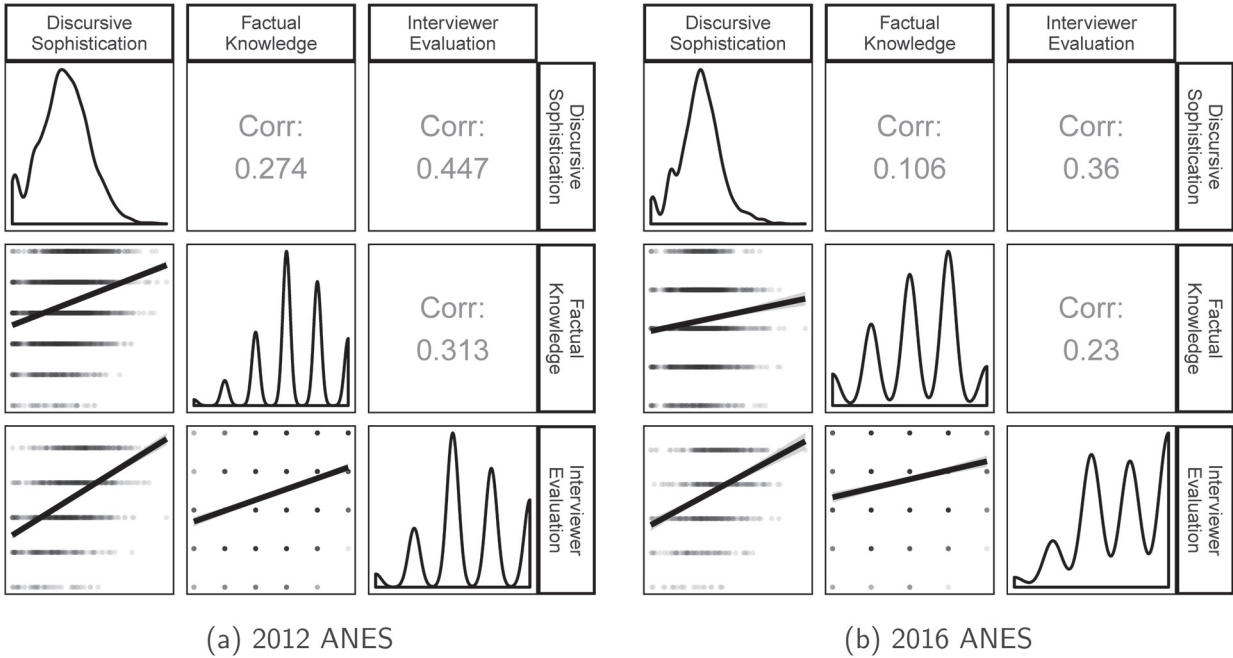


Figure 2.1: Correlation matrix of conventional political knowledge metrics and discursive sophistication. The plots on the diagonal display univariate densities for each variable. The panels in the lower triangular display the scatter plot of two measures as well as a linear fit. The upper triangular displays the correlation coefficient. All correlations reported are statistically significant with  $p < .05$ .

Figure 2.1 compares discursive sophistication to the conventional knowledge metrics for both surveys. Each figure presents scatterplots between individual measures (lower triangular), univariate densities (diagonal), and correlation coefficients (upper triangular). The measure of discursive sophistication is positively correlated with both conventional metrics while capturing some additional variation. Interestingly, there is a stronger correlation between discursive sophistication and interviewer evaluations than between factual knowledge and interviewer evaluations ( $r = .45$  vs.  $r = .31$  in 2012, and  $r = .36$  vs.  $r = .23$  in 2016). The open-ended measure therefore captures characteristics that influence subjective assessments of sophistication. Interviewers certainly form their impressions throughout the entire survey, but a respondent's verbatim answers seems to be more influential for subsequent knowledge assessments than a respondent's performance on the

<sup>7</sup>Interviewer assessments were only recorded in the face-to-face sample of the ANES.



## CHAPTER 2. LET'S TALK POLITICS

factual knowledge questions.

Overall, while discursive sophistication and the alternative measures are clearly correlated, the relationship between each metric is far from perfect. To provide some intuition as to whether the variation in discursive sophistication is theoretically meaningful, I present an example of open-ended responses of two individuals in the 2016 ANES who identified as Republicans and scored equally on the factual knowledge score (3 out of 4 correct responses), but varied highly in discursive sophistication. The results are presented in Table 2.1.

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

Table 2.1: Example of open-ended responses for low and high scores on discursive sophistication with equal factual knowledge scores (3 out of 4 correct responses). Column A displays the verbatim responses of an individual who scored low on discursive sophistication and column B displays the verbatim responses of an individual who scored high on the open-ended measure. Each row represents one of the likes/dislikes items included in the analysis. Note that the responses in this table were slightly redacted for readability (spelling errors removed, etc.).

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

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

## 2.6 Discursive Sophistication and Political Competence

I validate the measure of discursive sophistication by directly examining its effects on individual competences to perform political tasks in modern democracies (c.f., Lupia, 2006, 2015). More specifically, I consider the potential role of political sophistication in promoting (1) engagement and participation in politics, (2) the ability to incorporate new information, (3) precise positioning of candidates running for election, and (4) well-justified policy preferences. In the following, each point will be addressed using one of the three data sets described above.

### 2.6.1 Engagement and Participation in Politics

Political sophistication is often argued to promote individual engagement and participation in politics. In fact, factual knowledge items have been validated in the past based on their strong relationship with outcomes such as turnout and other forms of participation (c.f., Lupia, 2015, 230–233). Figure 2.2 compares the effects of discursive sophistication and factual knowledge in the 2012 and 2016 ANES on four dependent variables related to political engagement: turnout, non-conventional participation, internal efficacy, and external efficacy. The model predicting turnout is estimated via logistic regression while the estimates for the three remaining dependent variables are based on OLS. Each model equation includes both sophistication measures while controlling for gender, education, income, age, race, church attendance, survey mode (face-to-

## CHAPTER 2. LET'S TALK POLITICS

face vs. online), as well as the Wordsum vocabulary score measuring verbal intelligence.

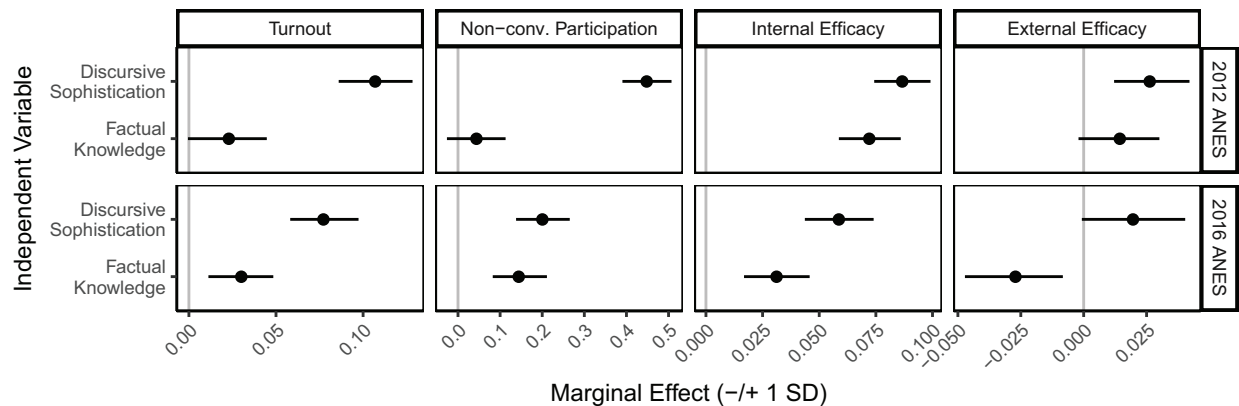


Figure 2.2: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2012 and 2016 ANES. For each dependent variable, the figure displays the change in expected values after increasing each sophistication measure from -1 to +1 standard deviation from its mean (including 95% confidence intervals). Model estimates are based on logistic regression (turnout) or OLS (non-conventional participation, internal efficacy, external efficacy). Both sophistication measures are included simultaneously while controlling for gender, education, income, age, race, church attendance, survey mode, and Wordsum vocabulary scores. Full model results are displayed in the appendix, Tables A.1 and A.2.

Each panel displays the expected difference in the respective dependent variable for a two standard deviation increase in each sophistication measure, while holding all other variables constant at their means. Overall, discursive sophistication is a stronger predictor of turnout, non-conventional participation, as well as (to a lesser extent) internal and external efficacy. In the 2012 ANES, the positive effect of factual knowledge on participation is statistically indistinguishable from zero when controlling for discursive sophistication. Furthermore, there is a negative effect of factual knowledge on external efficacy in the 2016 ANES. In contrast, the positive effect of discursive sophistication on external efficacy is more consistent with previous research. Considering these initial results, a potential concern may be that discursive sophistication is confounded by personality characteristics that influence verbatim response patterns as well as engagement. Appendix A.4 provides additional analyses controlling for such factors that might drive verbosity (extraversion and being reserved) as well as individual response length itself. The substantive conclusions remain unchanged.

### 2.6.2 Incorporation of New Information

Competent citizens should not only engage in politics but are also expected to be sufficiently informed about the issues of the day. As such, they have to be attentive to their media environments and incorporate potentially relevant new information about parties, office-holders, and policies. Indeed, Zaller (1990, 1992) and others argue that tests of factual information about politics are the best available proxy for awareness. In this analysis I draw on the 2015 YouGov study to explore whether discursive sophistication or factual knowledge serves as a better predictor of people’s ability to incorporate new information from media sources. As part of the survey, respondents were asked to read a newspaper article about a fictional infectious disease and were subsequently asked to answer questions about information provided in the article (e.g. regarding symptoms, modes of contraction etc.). I compute an additive index counting the pieces of information that were correctly recalled (*information retrieval*) as a measure of the ability to retrieve information from a news article on a non-partisan issue that is related to public health policies.

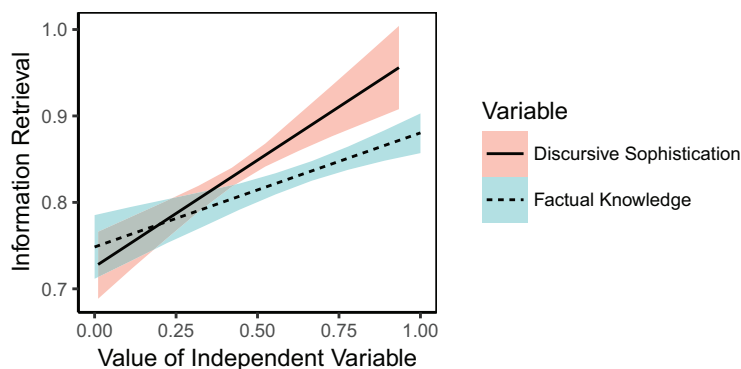


Figure 2.3: Expected information retrieval in the 2015 YouGov Study as a function of political sophistication (including 95% confidence intervals). Estimates are based on a linear regression model controlling for education, income, age, church attendance, gender, and race. Full model results are displayed in the appendix, Table A.3.

Figure 2.3 displays the relationship between political sophistication and disease information retrieval in the 2015 YouGov study. Estimates are based on a linear regression model controlling for education, income, age, church attendance, gender, and race. As a benchmark for discursive

sophistication, I again consider the effect of factual knowledge based on a battery of eight items similar to the knowledge questions in the ANES. Both discursive sophistication as well as factual knowledge increase the amount of information individuals are able to recall from a news article discussing a fictional disease. Similar to the previous results, the effects are stronger for discursive sophistication than for factual knowledge scores. The degree to which citizens discuss their own political beliefs in a more elaborate manner is not only a stronger predictor of political engagement but also serves as a better proxy for the ability to incorporate new information about a non-partisan issue.

### 2.6.3 Precise Positioning of Candidates

Citizens who are attentive to politics and able to incorporate new information should ultimately be better informed about the policies put forward by parties and political elites. This is a crucial component of citizen competence in representative democracies since precise knowledge about the policy positions held by candidates who are running for office allows voters to hold them accountable. Figure 2.4 presents the results of multiple heteroskedastic regressions where the error variance in candidate placements on multiple issues included in both ANES waves (general ideology, government spending, defense spending, health insurance policy, job guarantee, government assistance to Blacks, environment vs. jobs trade-off) is modeled as a function of discursive sophistication as well as factual knowledge (see Jacoby, 2006, for a similar procedure). More formally, each model for a given candidate placement on a specific policy issue takes the following form:

$$y \sim N(\mu, \sigma) \quad (2.5)$$

$$\mu = X\beta \quad (2.6)$$

$$\log(\sigma) = Z\gamma, \quad (2.7)$$

## 2.6. DISCURSIVE SOPHISTICATION AND POLITICAL COMPETENCE

where  $y$  is the vector of policy placements across respondents,  $X$  is a matrix of covariates predicting average candidate placements  $\mu$  (self-placement, education, income, age, church attendance, gender, race, and survey mode),  $Z$  denotes the covariates predicting the error variances  $\sigma$  (discursive sophistication, factual knowledge, Wordsum score), and  $\beta$  and  $\gamma$  are the parameters to be estimated.

The figure displays the estimated reduction in error variances of candidate placements when each sophistication measure is increased by two standard deviations. Larger negative values indicate a stronger reduction in error variances and hence more precise candidate placements. Both factual knowledge and discursive sophistication significantly decrease error variances in policy placements of presidential candidates. Some interesting differences, however, emerge when comparing both waves of the ANES. In the 2012 election, discursive sophistication in open-ended responses was a slightly weaker predictor of precise candidate placements than performance on factual knowledge quizzes across multiple issues. This picture is reversed in the 2016 election, where more elaborate open-ended responses were strongly predictive of precise candidate placements. Factual knowledge, on the other hand, did not significantly improve the precision of candidate placements for multiple issues.

### 2.6.4 Well-Justified Policy Preferences

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

## CHAPTER 2. LET'S TALK POLITICS

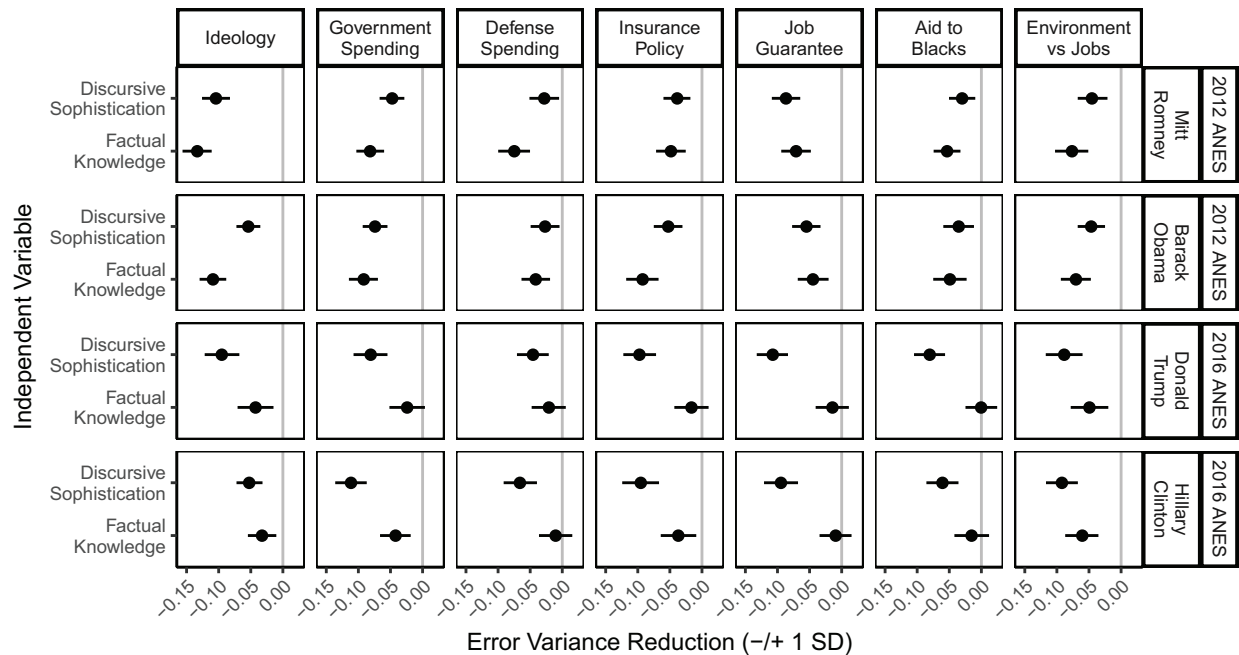


Figure 2.4: Error variance reduction in candidate placements on multiple issues in the 2012 and 2016 ANES. The figure displays the difference in estimated error variances after increasing each sophistication measure from -1 to +1 standard deviation from its mean (including 95% credible intervals). Models are estimated in Stan using non-informative priors. Detailed model results including convergence statistics are displayed in the appendix, Tables A.4 and A.5.

The results are presented in Figure 2.5. Since the Swiss post-referendum surveys were conducted in three different languages (German, French, and Italian), I computed the measure of discursive sophistication separately for each group of respondents. The figure displays the distribution of discursive sophistication for each level of justification coded by Colombo (2016) as well as the correlation coefficients for both respective variables. Across all three language groups, discursive sophistication is systematically higher among respondents with the highest level of justification and both measures are positively correlated ( $r = 0.29, 0.25, \text{ and } 0.35$ , respectively). The proposed measure of discursive sophistication therefore shows a high degree of correspondence with individual levels of justification assessed by independent manual coders.

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

## 2.6. DISCURSIVE SOPHISTICATION AND POLITICAL COMPETENCE

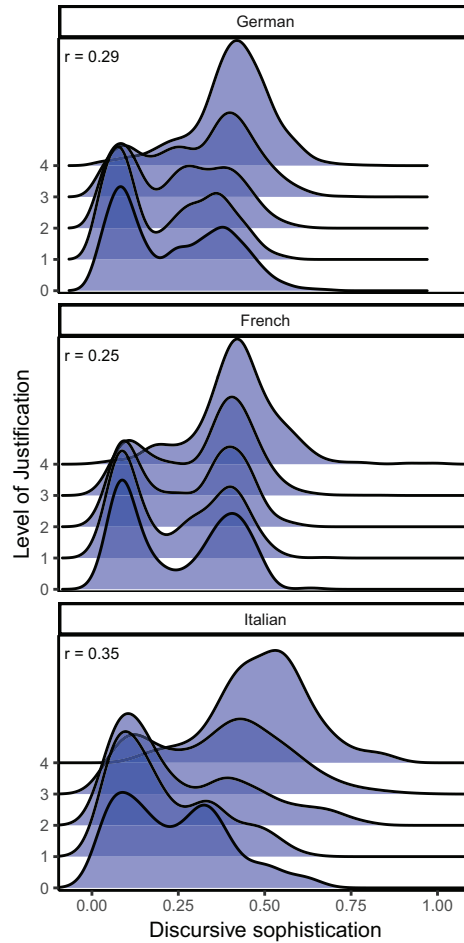


Figure 2.5: Discursive sophistication and manually coded level of justification (Colombo, 2016) in Swiss post-referendum surveys. The plot compares kernel densities of discursive sophistication for each manually coded level of justification.

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



### 2.7 Application: The Gender Gap in Political Knowledge

A common finding in public opinion research is the fact that women have lower levels of observed political knowledge than men. For example, Verba, Burns, and Schlozman (1997) report that women score lower on political information, interest, and efficacy, which decreases their respective levels of political participation. Since gender differences in political information and interest can only partly be explained by resource-related factors such as individual levels of education, the authors diagnose a “genuine difference in the taste for politics” between men and women, which they suspect to be driven largely by socialization (see also Wolak and McDevitt, 2011). Indeed, Dow (2009, 117) describes the systematic gender differences in knowledge “one of the most robust findings in the study of political behavior.”

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

## 2.7. APPLICATION: THE GENDER GAP IN POLITICAL KNOWLEDGE

Overall, the gender gap has been shown to be influenced by how we ask for political information in surveys, as well as the kind of knowledge that is required for a correct response. Indeed, a comprehensive cross-national analysis of election studies in 47 countries between 1996 and 2011 suggests that question format and content account for large portions of the variance of gender disparities in political knowledge (Fortin-Rittberger, 2016).

### 2.7.1 Descriptive Results

How do men and women compare on the different metrics of political sophistication in the surveys analyzed in the present study? Figure 2.6 displays the average levels of discursive sophistication as well as conventional metrics comparing both genders. While we observe a sizable and statistically significant gender gap for factual knowledge in both ANES surveys, this difference disappears for discursive sophistication. These results are replicated in the 2015 YouGov survey. As before, we observe a significant gender gap in factual knowledge which disappears using the discursive measure. Of course, it is important to ask whether this absence of a gender gap in discursive sophistication is theoretically meaningful or rather an artifact of the measurement approach itself.

One way to investigate this question is to explore gender differences in discursive sophistication using the Colombo (2016) data and comparing them to her manually coded measure. That way, we can not only examine whether the lack of a gender gap in discursive sophistication replicates using an additional survey, but also check whether there is an equivalent lack of gender differences in Colombo's alternative measure of citizen competence in direct democracies. If discursive sophistication captures a person's motivation to undertake in-depth processing and form quality opinions (and assuming these characteristics do not differ by gender), there should be no difference between men and women on either metric (discursive sophistication and Colombo's measure). As shown in the bottom row of Figure 2.6 there are indeed no significant gender differences on *both* metrics across all three languages in the Swiss referendum surveys. The absence of a gender gap is consistent whether open-ended responses are coded manually or using the proposed measure

of discursive sophistication.

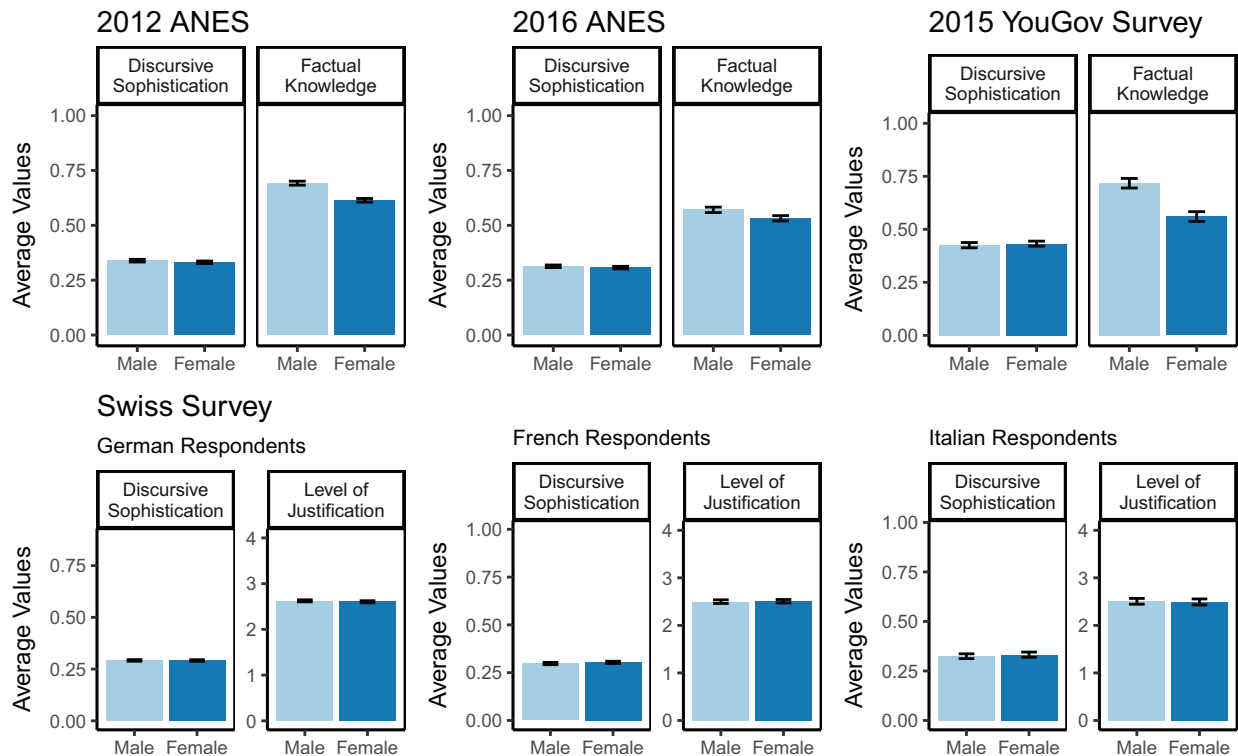


Figure 2.6: The gender gap in political sophistication. The figures display mean levels of sophistication for each measure comparing men and women (including 95% confidence intervals). Gender differences in factual knowledge in the 2012/2016 ANES and 2015 YouGov survey (top row) are statistically significant with  $p < .05$ . Gender differences in discursive sophistication and manually coded levels of justification (Colombo, 2016) are not statistically significant.

## 2.7.2 Controlling for Alternative Explanations

Prior research suggests that at least part of the gender gap in political knowledge can be attributed to real discrepancies in resources and engagement. To the extent that differences between men and women can be explained by these underlying factors, they are less likely to be an artifact of the measurement of knowledge itself. Accordingly, we need to control for determinants of political knowledge to provide a more comprehensive examination of the veracity of observed gender differences. Figure 2.7 displays estimated effects of various potential common determinants of factual knowledge and discursive sophistication on both measures. Previous studies consistently

## 2.7. APPLICATION: THE GENDER GAP IN POLITICAL KNOWLEDGE

showed that political information levels are positively related to high media exposure, frequent political discussions, education, and income. Furthermore, I include age, race, church attendance, and survey mode (face-to-face vs. online) as additional control variables.

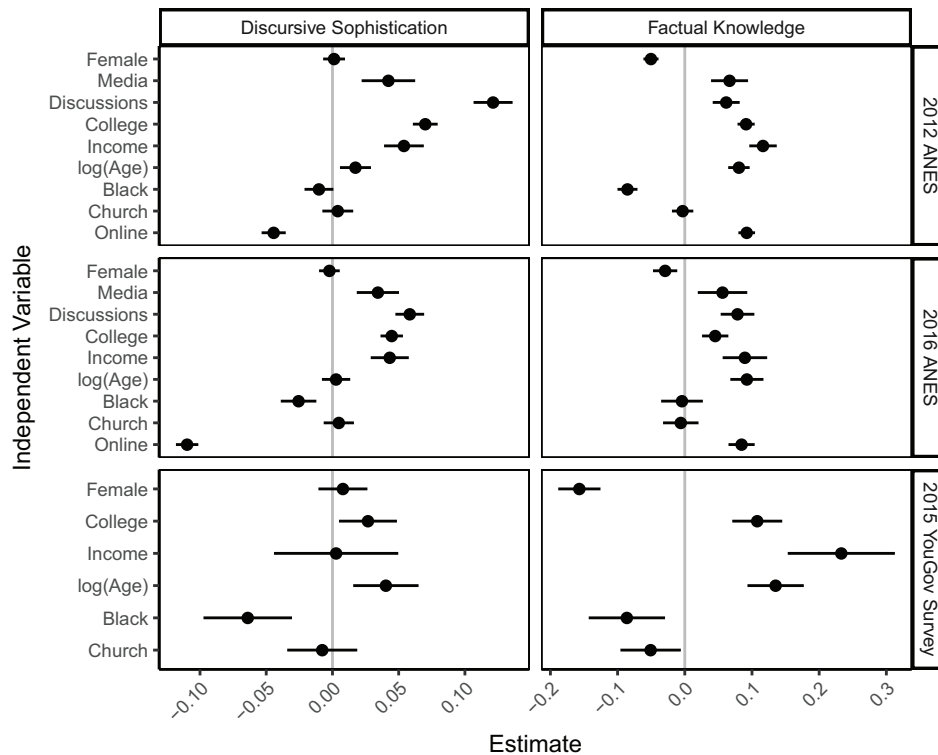


Figure 2.7: Common determinants of political sophistication. Estimates are OLS regression coefficients with 95% confidence intervals. Dependent variables are discursive sophistication as well as conventional metrics of political knowledge. Full model results are displayed in the appendix, Tables A.6 and A.7.

After controlling for common determinants, discursive sophistication again reveals no significant differences between men and women in both ANES surveys as well as the 2015 YouGov study. The gender gap in factual political knowledge, however, persists and is substantively as well as statistically significant after controlling for various resource-related factors. Even though women do not perform as well as men on political quizzes, they do not differ substantially in complexity and sophistication when they describe their political preferences. The effects of the remaining variables are quite similar across both measures and different surveys. Knowledge and sophistication is significantly higher among respondents who are more exposed to political news

media, discuss politics frequently, are more educated, and have higher income.<sup>8</sup> Overall, the finding that determinants of political sophistication are consistent across models lends additional validity to the open-ended measure.

To summarize, we only observe a significant gender gap when looking at conventional recall-based measures, a result that previous research (at least partly) attributed to the content (i.e., focusing on issues that are less relevant to women) and format (i.e., stereotype-threat and guessing) of the question batteries. When using the alternative measure—discursive sophistication—any evidence for systematic differences between men and women disappears.

### 2.7.3 Explaining the (Lack of the) Gender Gap

If it is the case that women are able to close the gender gap in discursive sophistication because they are able to focus on different considerations that are salient to them when discussing their political preferences, we should observe systematic variation in the issues men and women discuss in open-ended responses. Based on the structural topic model used to compute discursive sophistication, I now examine the subset of topics that showed the largest absolute gender difference in topic prevalence in the 2012 and 2016 ANES. The results are displayed in Figure 2.8.

Positive coefficients indicate that women are more likely than men to mention a given topic, and vice versa. As such, the top six topics are more prevalent among men and the bottom six have a higher probability to be mentioned by women. Each coefficient is labeled with the five highest probability terms related to the topic to illustrate its content. Across both ANES studies, women were less likely than men to discuss foreign affairs, economic issues, or the Supreme Court. Instead, they focused on issues related to women's rights, equality, or health care. The considerations taken into account by women when discussing their political preferences

---

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

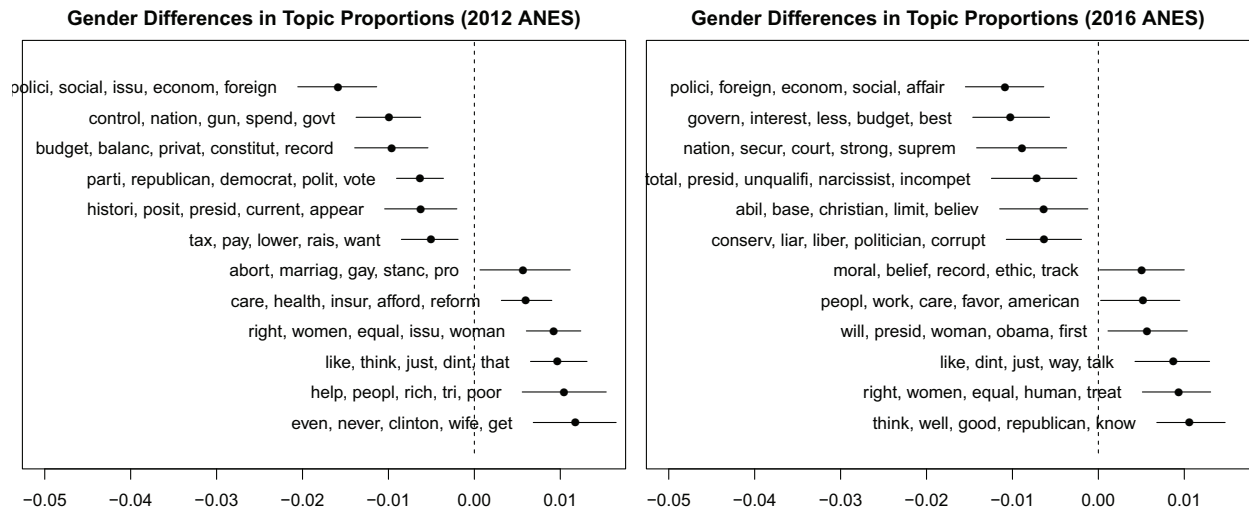


Figure 2.8: Gender differences in topic proportions in open-ended responses based on the structural topic model used to compute discursive sophistication (including 95% confidence intervals). Coefficients indicate the difference in predicted topic prevalence among men and women; positive values indicate higher prevalence among women. Labels are based on the five highest probability terms related to the topic.

are therefore clearly different from men's and—crucially—the issues raised by men happen to be more aligned with what political scientists often deem as necessary information (i.e., pertaining to the economy, institutions, elites, etc.). Yet, from a normative perspective, there is no reason to assume that one set of issues should be more important for citizens when forming their political preferences and making competent voting decisions.

## 2.8 Conclusion

Political scientists should worry less about pure levels of *factual knowledge* and instead focus on the necessary conditions for individuals to make *competent* decisions. Competence in the context of political decision-making and voting requires citizens to hold informed attitudes about their representatives. Factual knowledge about political institutions might be a useful proxy for competence in certain scenarios. However, it cannot address directly whether individuals hold well-considered opinions about political actors they try to hold accountable. In comparison, the

measure of discursive sophistication proposed here is agnostic about the specific contents of people's beliefs, but directly captures the complexity of individual attitude expression. Furthermore, it can be easily applied to assess sophistication in any decision-making context (such as a policy referendum or a local election) by fielding targeted open-ended questions related to the relevant underlying beliefs and preferences.

The findings presented in this paper show that conventional knowledge indices and the discursive measure share a substantial amount of variance. However, they are far from being identical and capture different aspects of sophistication. Most importantly, using the discursive measure, evidence for the gender gap commonly reported using factual knowledge scales disappears. Women might know fewer facts about political institutions, but they do not differ substantively in the complexity of their expressed political beliefs. The fact that women perform just as well as men on discursive sophistication across various surveys can be attributed to the fact that they focus on different considerations when evaluating political parties and candidates. This issue has long been recognized in the literature (e.g., Graber, 2001; Dolan, 2011), but it cannot be properly addressed while relying exclusively on off-the-shelf political knowledge batteries. As discussed at the outset, Zaller (1992) and others made the argument that testing for factual information provides the best measure of political awareness as it captures "what has actually gotten into people's minds, which, in turn, is critical for intellectual engagement with politics" (21). The results presented in this paper suggest that a direct examination of open-ended responses provides a viable alternative approach.

## Is It All About Values?

### Measuring Morality in Political Attitude Expression\*

#### Abstract

This study explores whether and how individuals evoke moral considerations when discussing their political beliefs. Analyzing open-ended responses in the 2012 American National Election Study (ANES) using a previously validated dictionary, I find systematic ideological differences in moral reasoning—even when respondents are not explicitly asked about morality. The study proceeds to show that the reliance on moral considerations in attitude expression is amplified by the moral content of individual media environments.

---

\*This chapter is a reprint of a short article: Kraft, Patrick W. 2018. "Measuring Morality in Political Attitude Expression." *The Journal of Politics* (forthcoming): <https://doi.org/10.1086/696862>. The code for the chapter is available on GitHub: <https://github.com/pwkraft/mft>.



### 3.1 Introduction

Increasing levels of polarization have renewed scholarly interest in the psychological and attitudinal differences between liberals and conservatives (Jost, 2006). One such area of research focuses on the moral underpinnings of ideology. According to *Moral Foundations Theory* (MFT), moral thinking is organized by at least five dimensions: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation (Graham et al., 2013). Liberals and conservatives differ in their emphasis on each foundation, with liberals prioritizing care and fairness, and conservatives endorsing all five dimensions equally (Graham, Haidt, and Nosek, 2009).

A series of recent studies shows that the moral foundations influence issue preferences (Kertzer et al., 2014), candidate trait evaluations (Clifford, 2014), and vote choice (Iyer et al., 2010). Research further suggests that moral framing in elite communication can elicit attitude change (e.g. Clifford et al., 2015; Feinberg and Willer, 2013). For the most part these studies measure moral reasoning with the Moral Foundations Questionnaire (MFQ), which explicitly asks respondents to judge the importance of considerations related to the five foundations (e.g., Graham et al., 2011). Yet, by explicitly asking about morality, researchers presuppose an important link that requires more careful empirical investigation.

The present study explores how people utilize moral arguments in day-to-day political reasoning in a more unobtrusive context. Using a moral dictionary validated in previous studies (Graham, Haidt, and Nosek, 2009), I propose a novel approach to analyze individual verbatim responses to open-ended likes/dislikes questions in the 2012 American National Election Study (ANES). Measuring moral reasoning in open-ended responses directly captures whether political attitudes are infused by morality without being prompted by the language of a questionnaire. Insofar as moral intuitions play a role in political attitude expression, citizens should rely on the moral foundations when discussing their opinions about political actors, even if not explicitly asked to do so.

The analysis begins by replicating previous findings regarding MFT and ideology using the

open-ended measure. Consistent with MFT, the results reveal systematic differences between liberals and conservatives in the reliance on specific moral considerations. Furthermore, these differences in verbatim moral reasoning predict candidate preferences and vote choice—even after controlling for a person’s party identification. Integrating a large-scale content analysis of individual media environments, I proceed to show that people who are exposed to moral rhetoric in political news are more likely to rely on moral considerations when discussing their political beliefs. Overall, this study improves conventional dictionary-based approaches to analyze open-ended responses and showcases the integration of individual media environments to trace the influence of media exposure on attitude expression.

## 3.2 Method

This study utilizes the moral foundations dictionary created by Graham, Haidt, and Nosek (2009) to identify references to specific moral considerations when respondents discuss what they like and dislike about political parties and candidates.<sup>1</sup> Other studies have used (variations of) this dictionary to identify the moral foundations in elite communication (e.g., Clifford et al., 2015) or political advertising (e.g., Lipsitz, 2018), but to date no research has examined verbatim attitude expressions in surveys. Based on the terms signaling each foundation in the dictionary, any document can be scored according to its emphasis on the respective moral dimension. Conventional dictionary-based methods usually consist of the proportion of signal word occurrences in each document (e.g., Graham, Haidt, and Nosek, 2009). However, some dictionary terms are problematic when applied to verbatim survey responses. In particular, certain words might be too ubiquitous to be regarded as an unambiguous indicator for specific moral considerations. For example, “leader” is a signal word for the authority dimension. However, respondents may describe the qualities of presidential candidates as *leaders* irrespective of moral considerations related to authority.

<sup>1</sup>See Appendix B.1 for the full content of the dictionary.

One way to address this problem would be to revise the dictionary and eliminate ambiguous words. Yet such revisions could be arbitrary and leave too much discretion to the researcher. Drawing on techniques developed in the field of information retrieval, I propose an alternative approach. If a specific dictionary term like “leader” is commonly used to describe presidential candidates, it is likely that the term can be used in multiple contexts and is not necessarily unique to the moral domain. Terms that are used by almost all respondents therefore provide less information about differences in their (moral) reasoning than terms that only occur in few responses. In this study, *MFT scores* are computed for a foundation by weighting each term in the dictionary according to its ubiquity across documents, which serves as a proxy for the term’s discriminative information:

$$\text{MFT}_{if} = \frac{1}{W_i} \sum_{t \in \mathcal{D}_f} \left[ w_{it} * \log_{10} \left( \frac{N}{n_t} \right) \right], \quad (3.1)$$

where  $\text{MFT}_{if}$  denotes the score of document  $i$  for foundation  $f$ ,  $W_i$  is the total number of words in document  $i$ ,  $t$  indicates a term in the set of signal terms in foundation dictionary  $\mathcal{D}_f$ ,  $w_{it}$  denotes the number of occurrences of term  $t$  in document  $i$ ,  $N$  represents the total number of documents, and  $n_t$  is the number of documents in which the term  $t$  appears. The weight represents the inverse of the proportion of documents in which the target term appears.<sup>2</sup> Terms that are ubiquitous across the entire corpus receive a lower weight, and terms that appear in only few documents receive a higher weight.

Each document is an individual’s verbatim response to a set of open-ended questions. As such, a respondent’s MFT score for foundation  $f$  is the weighted proportion of words in the response that signal the respective foundation. The score has a lower bound of 0 (document does not contain any dictionary terms) and is independent of document length (since it is based on relative occurrences). Higher scores imply larger proportions of dictionary terms in a document. Most importantly, however, words that appear in nearly all open-ended remarks affect MFT scores less

<sup>2</sup>This specification is usually referred to as tf-idf weighting and is commonly used in quantitative text analysis (see Manning et al. 2008, ch. 6 for an introduction).

than words that appear only in a few responses because ubiquitous terms convey less information about differences across individuals. Overall, the MFT score provides a correction for potential distortions due to suboptimal terms in the dictionary. Since nominal values of the MFT score above zero do not have a clear substantive interpretation, they are rescaled to unit variance.

## 3.3 Results

Open-ended responses to the likes/dislikes items in the 2012 ANES were aggregated for each individual and pre-processed by correcting spelling errors before computing MFT scores. Appendix B.2 provides detailed information on the procedure, including the weighting scheme and raw proportions of individuals mentioning each foundation. Results for the sanctity dimension are not presented below due to its low general prevalence in individual attitude expressions.<sup>3</sup> To account for confounding factors related to the respondents' eloquence when discussing their political attitudes, all models reported below include controls for education, logged overall response length, as well as the Wordsum vocabulary score measuring verbal intelligence.

### 3.3.1 Ideological Differences

MFT scores measure the weighted proportion of moral foundation terms in an open-ended response. Since they are bounded at zero (i.e., response does not contain any moral words), I begin by estimating a set of Tobit regressions using ideology to predict individual MFT scores for each moral foundation.<sup>4</sup> Figure 3.1 compares liberals and conservatives while holding all other variables constant at their respective means. To facilitate their substantive interpretation, I decompose the estimates into the effect of ideology on the probability of mentioning a specific foundation *at all* (i.e., the probability the MFT score is larger than zero) as well as the degree of emphasis on the foundation given that it was mentioned by a respondent (i.e., the change in the MFT score

<sup>3</sup>Only about 3.6% of respondents mentioned the sanctity dimension.

<sup>4</sup>Full estimates for this and all subsequent models are presented in Appendix B.4.

given that it is larger than zero, measured in standard deviations).<sup>5</sup>

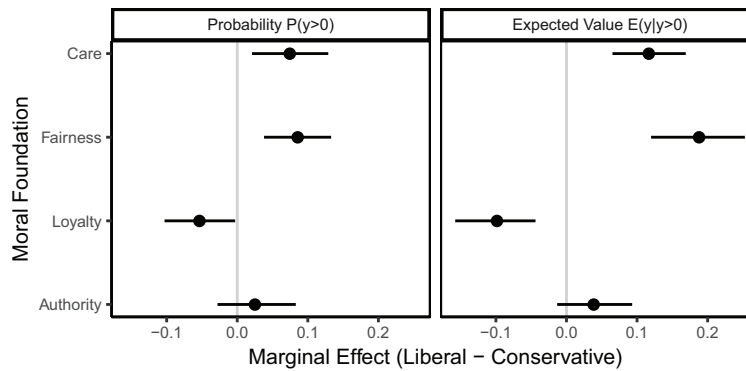


Figure 3.1: Difference between liberals and conservatives in the probability of mentioning a moral foundation (left panel) and in the MFT score given that the foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables include age, sex, race, church attendance, survey mode, education, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix, Table B.3.

Positive values denote a higher probability of mentioning the respective moral foundation (left panel) or a higher MFT score (right panel) among individuals who identified as liberals, while negative values indicate a higher probability/higher score among conservatives. The effects are consistent with the expectations of MFT for three out of four moral foundations. Liberals are about 8 percentage points more likely than conservatives to mention the foundations of care and fairness. Furthermore, given that respondents mention these two foundations at all, liberals emphasize it more than conservatives when evaluating political parties and candidates. The MFT score for the care foundation is about 0.12 standard deviations higher among liberals than conservatives. The effect is slightly larger for the fairness dimension. Conversely, being conservative is associated with an increased loyalty MFT score by about 0.1 standard deviations. There are no significant differences between liberals and conservative on the authority dimension.

<sup>5</sup>See for example McDonald and Moffitt (1980) for details on decomposing Tobit estimates.

### 3.3.2 Moral Considerations and Vote Choice

A skeptic may worry that the verbal expression of moral considerations might not be as strongly related to other forms of political behavior (e.g., vote choice) as moral foundations measured by the MFQ. To address this concern, Figure 3.2 presents the changes in expected probabilities of voting for the Democratic (vs. Republican) presidential candidate in the 2012 election for individuals emphasizing the moral foundations in their open-ended responses. The estimated probabilities are based on logit models including MFT scores for each moral foundation as independent variables as well as controls for various sociodemographic characteristics. Individuals who emphasized moral considerations related to the care and fairness foundations were more likely to vote for Barack Obama than for Mitt Romney. Respondents who emphasized the loyalty foundation, on the other hand, were less likely to vote for Obama.<sup>6</sup>

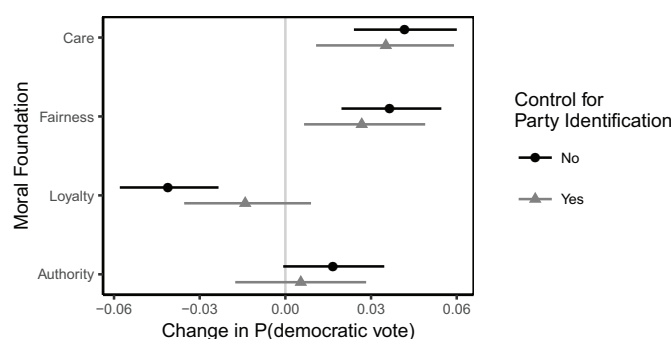


Figure 3.2: Change in predicted probabilities of voting for the Democratic rather than Republican candidate when MFT score is increased from its minimum (no overlap between dictionary and response) by one standard deviation, holding control variables constant at their respective means (along with 95% confidence intervals). Control variables include party identification, age, sex, race, church attendance, survey mode, education, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix, Table B.4

The effects on vote choice might not seem large, but bear in mind that the measure of moral reasoning is based solely on the content of open-ended responses in which respondents were *not* explicitly asked about morality. The fact that moral considerations evoked by respondents are nevertheless related to their political preferences indicates that their open-ended comments

<sup>6</sup>Appendix B.3.3 shows similar results in an analysis of feeling thermometers towards parties and candidates.

about both candidates and parties are imbued with moral content that in turn relates to political judgments in the manner suggested by MFT. As such, analyzing *how* individuals talk about their political preferences prior to an election helps us predict their subsequent vote choice.

### 3.3.3 Media Content and Exposure to Moral Rhetoric

It can also be informative to examine the reliance on moral considerations *in general* rather than focusing only on individual foundations. For example, a recent study found that moral language in political ads elicits emotional responses among recipients (Lipsitz, 2018). Here, I investigate whether exposure to moralized discourse in the media is associated with a stronger general reliance on moral considerations in attitude expression. For each individual, I compute the sum of MFT scores to measure emphasis of *any* moral foundation. The main independent variable captures moralization of media environments based on a content analysis of media sources consumed by each individual. Using Lexis-Nexis, I retrieved the content of 28 media sources covering either presidential candidate during the survey field period in the last month of the campaign (October 2012) and coded the emphasis on moral considerations using the weighted dictionary approach described earlier.<sup>7</sup> Based on each source's content, I create a measure that represents the extent to which each individual's media environment emphasized moral considerations by averaging (median-centered) MFT scores of all media outlets retrieved by a respondent.

Figure 3.3 presents the results of a Tobit model where effects are again decomposed into the probability of mentioning any moral foundation (left panel) as well as the emphasis on morality, given that any foundation was mentioned (right panel). Individuals who are exposed to media sources that report on the campaign in a more moralized manner put a stronger emphasis on moral considerations in their open-ended responses. Although it is difficult to determine the causal ordering, the pattern is consistent with research suggesting that people adopt moral arguments from their media environment (e.g., Clifford et al., 2015).

---

<sup>7</sup>Sources include e.g., New York Times, CNN.com, and Fox News Programs. See Appendix B.2.2 for details.

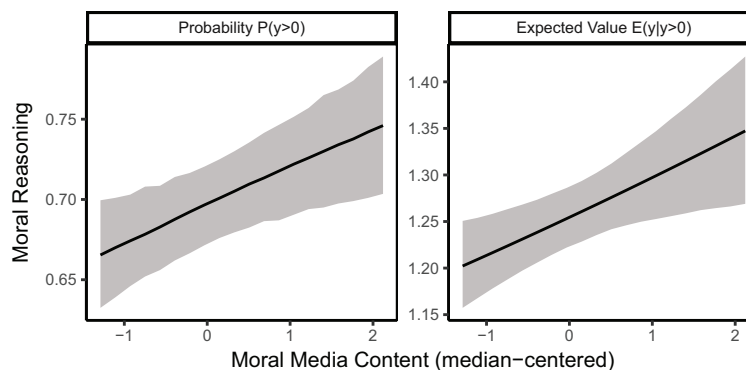


Figure 3.3: Effect of MFT content in individual media environments on the probability of mentioning any moral foundation (left panel), and on the summed MFT score given that any foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables include political knowledge, general media exposure, political discussion frequency, age, sex, race, church attendance, survey mode, education, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix, Table B.5.

### 3.3.4 Robustness Checks

To this point, the analyses assume that the dictionary-based approach captures the theoretical concept of interest—*morality*. Yet, the terms in the dictionary may also be recovering other (i.e., non-moral) patterns in word choice. Appendix B.3 presents the results of multiple supplementary analyses to alleviate this concern for both open-ended responses as well as media environments. Two results are briefly highlighted here.

First, could the ideological differences in open-ended responses be explained by the survey context? To test for this possibility, I replicate the analysis from Figure 3.1 using data from a random-digit-dial adult sample of residents within a 25 mile radius of a large northeastern state university conducted between early January, 2001 and July, 2003. Compared to the ANES, the survey varied the mode (phone), political context (non-election year, Republican presidency), as well as the set of open-ended items (discussing liberals and conservatives as social groups rather than candidates and parties). Notwithstanding these changes, the ideological differences in moral reasoning are consistent with the results presented above.



Second, even if the patterns can be replicated, can the dictionary really capture underlying moral rhetoric? In an additional analysis, I compare dictionary-based MFT scores with individual assessments of moralization conducted by an independent group of researchers. Feinberg and Willer (2013) explored moral rhetoric in a set of 232 newspaper op-eds on environmental issues by asking a group of coders to assess the degree to which they used rhetoric grounded in moral foundations. Moralization measured using MFT scores is positively correlated with individual coder assessments that did not rely on any dictionary (although modestly,  $r = 0.27$ ).

### 3.4 Discussion

Moral Foundations Theory has become an influential framework for understanding ideology and political attitudes. Yet, existing measures fail to directly assess whether individuals rely on moral considerations in their day-to-day political reasoning. I address this gap by examining moral arguments in individual attitude expression. Consistent with MFT, there are systematic patterns in the emphasis on moral considerations among liberals and conservatives for three out of four foundations. Liberals are more likely to mention considerations related to care and fairness, whereas conservatives are more likely to emphasize the moral foundation of loyalty. Moreover, morality in attitude expression is related to vote choice and the exposure to moralized political discourse in the mass media is associated with increased reliance on moral considerations.

That said, ideological differences on binding foundations (loyalty, authority) appear less persistent than those on individualizing foundations (care, fairness). Appendix B.3.2 examines potential explanations by analyzing subsets of the dictionary and the open-ended items. Liberals are more likely to mention the authority foundation in the context of positive moral endorsements (*virtues*) when they discuss aspects they *liked* about their *in-party*. In contrast, there is suggestive evidence that conservatives are more likely to discuss the authority dimension in the context of negative endorsements (*vices*). This finding indicates that liberals and conservatives attach diverging meanings to certain foundations, which promises to be a fruitful area for future research.

Overall, this study improved conventional dictionary-based approaches in order to utilize a largely neglected data source: verbatim open-ended responses. Using this method, scholars can study moral reasoning in surveys that do not contain the MFQ simply by relying on open-ended items. Lastly, the approach outlined here allows for a seamless integration of media content in the analysis of moral reasoning, which can further illuminate how exposure to political discourse fosters ideological differences in moral reasoning. In times of growing partisan polarization, a better understanding of the antecedents of this ideological divide is essential.

# Chapter 4

## Change My View:

### Do Moral Appeals Facilitate Compromise?\*

#### Abstract

The American electorate is becoming increasingly polarized. According to research in moral psychology, these growing disagreements between liberals and conservatives can be attributed to fundamental differences in the moral frameworks that shape individual ideology. Indeed, scholars suggest that ideologues would be more likely to reach compromise if both sides spoke the same “moral language.” While this implicit assumption has intuitive appeal, it remains largely untested empirically. Drawing on a unique dataset from the online discussion board *Reddit*, this paper examines how moral appeals can affect individual persuasion and the likelihood of compromise.

\*The code for this chapter is available on GitHub: <https://github.com/pwkraft/cmV>.

## 4.1 Introduction

Recent years have witnessed a resurgence in partisan polarization in the United States. Politically engaged citizens hold more diverging policy views, are more ideologically extreme, and exhibit stronger negative affect towards out-partisans than in the past (Hetherington, 2001; Abramowitz and Saunders, 2008; Iyengar, Sood, and Lelkes, 2012; Mason, 2015; Huddy, Mason, and Aarøe, 2015; Iyengar and Westwood, 2015). A growing literature in moral psychology attributes this divide (at least partially) to fundamental differences in moral frameworks that guide liberal and conservative thinking (e.g., Haidt, 2012; Graham et al., 2013). A recent analysis by Garrett and Bankert (2018), for example, finds that individual tendencies to moralize politics exacerbates affective polarization between Democrats and Republicans, which ultimately results in greater social distance and hostility towards out-partisans. More generally, moral conviction as an attribute of attitude strength has been shown to have wide-ranging behavioral consequences (Skitka, Bauman, and Sargis, 2005; Skitka and Morgan, 2014), including diminishing people's willingness to compromise in the realm of politics (Ryan, 2014, 2017).

Do these findings imply that morality in politics is always bound to foster disagreements and hostility between opposing views? Recent research building on Moral Foundations Theory pioneered by Haidt (2007) and colleagues suggests otherwise. According to this view, disagreements about morality are rooted in the underlying intuitions that form people's moral frameworks (Haidt, 2012). For instance, differential emphasis on basic moral dimensions predicts attitudes on culturally divisive issues such as abortion, the death penalty, or same-sex marriage (Koleva et al., 2012). More importantly, however, speaking the same "moral language" can overcome ideological divides. Indeed, political arguments can persuade individuals holding opposing views to the extent that they are emphasizing common moral ground (e.g., Day et al., 2014; Feinberg and Willer, 2015). Moral frames that rely on this logic, for example, were shown to be effective in convincing conservatives to embrace environmental protection policies and sustainable behavior (Kidwell, Farmer, and Hardesty, 2013; Feinberg and Willer, 2013).

However, few studies examined the persuasiveness of congruent moral appeals beyond the context of simple framing scenarios. Instead, they mostly focus on the effect of isolated messages without giving participants real opportunities to respond or engage in a dialogue. Political discourse is more complex and it is therefore unclear whether previous findings directly translate into more dynamic environments. Accordingly, the suggested potential of moral arguments to help overcoming disagreements—for example in the context of political discussions—is largely assumed as a potential implication and has not been subjected to a direct empirical test. Political discussions are an important source of information for citizens (Huckfeldt et al., 1995) and they have been shown to increase engagement and tolerance of opposing views (Mutz, 2002). Furthermore, Druckman and Nelson (2003) demonstrate that elite framing effects—often viewed as a potential source of polarization—can be mitigated by discussions in heterogeneous groups. Other research shows that such conversations can overcome polarization and partisanship (Klar, 2014). Notwithstanding, most research on deliberation pays little attention to the actual discussion contents (see Barabas, 2004; Karpowitz, Mendelberg, and Shaker, 2012; Mendelberg, Karpowitz, and Oliphant, 2014, for notable exceptions). As a result, we know very little about the role of moral arguments as a potential moderator of discussion effects, which—depending on the perspective in moral psychology—might hurt or harm the potential for compromise.

The present study fills this gap by analyzing the content of more than 10,000 conversations on the active *Reddit* community /r/ChangeMyView<sup>1</sup> (CMV). Discussions on CMV—which are anonymous but at the same time successful in maintaining civil discourse—provide an ideal environment to explore correlates of argument persuasiveness across a wide array of topics. For the analyses presented here, I rely on a dataset of matched argument pairs extracted from CMV by Tan et al. (2016), who focused on the role of linguistic features that predict argument strength. My analysis extends these results by examining the effects of moral appeals on attitude change. The findings show that moral arguments can facilitate compromise, but only to the extent that they are congruent with the moral framework of the opposing discussant.

---

<sup>1</sup><https://www.reddit.com/r/changemyview/>

## 4.2 Theoretical Background

Politics is centered around persuasion and the exchange of opposing arguments. Officeholders, legislators, and activists spend much of their time trying to convince citizens to support one policy over another. As Cobb and Kuklinski (1997) eloquently note, “[p]ersuasion, changing another’s beliefs and attitudes, is about influence; and influence is the essence of politics” (88-89). Of course, attempts to persuade are not only limited to elite communications. Citizens discuss political issues with their peers, which turns social networks into a major information source influencing individual attitudes (e.g., Huckfeldt et al., 1995; Ahn, Huckfeldt, and Ryan, 2010; Lazer et al., 2010). The following sections briefly discuss previous approaches to persuasion in politics and connects them to research in moral psychology that helps inform our understanding of the nature of compelling arguments.

### 4.2.1 Two Routes to Persuasion

One influential framework to conceptualize and explain persuasive communication is the Elaboration - Likelihood Model (ELM) developed by Petty and Cacioppo (1986a,b). The theory distinguishes two separate routes to persuasion, each characterized by their distinctive consequences for a message’s effectiveness to change people’s attitudes. The first type—the *central route*—is a result of thoughtful processing and a thorough evaluation of the argument’s merit. According to this process, people who are sufficiently motivated will incorporate arguments after careful consideration and update their attitudes accordingly. The second type of persuasion, on the other hand, does not require elaborate processing but rather relies on simple cues based on the source of the argument (e.g., group membership, attractiveness, etc.). This route to persuasion is called the *peripheral route* and it can operate without much scrutiny regarding the content of the message (see also Chaiken and Eagly, 1989, for a similar distinction between systematic and heuristic processing). It follows from this distinction that people’s motivation and capability to engage in elaborate processing determines whether the persuasiveness of communications is

driven by argument strength itself or rather peripheral cues.

Since contextual factors and individual predispositions affect whether messages are closely scrutinized, different types of arguments may be more or less effective under varying circumstances. For example, Cobb and Kuklinski (1997) analyze the influence of an argument's complexity on its persuasiveness in two issue areas (NAFTA and health care). Interestingly, they find that while complex arguments were more compelling in the context of international trade, simple arguments proved more effective when discussing the issue of health care. However, the question of why these differences arise is left largely unanswered by Cobb and Kuklinski (1997). One explanation for the inconsistencies is the variation in people's motivation and ability to engage in more thoughtful processing (i.e., their elaboration likelihood). In the absence of such motivation, they are more likely to rely on peripheral cues which renders complex arguments ineffective. A potential motivating stimulus may be the argument's linkage to a person's values. For example, Nelson and Garst (2005) presents experimental evidence showing that people are paying more attention to messages that are consistent with their own value orientation. Participants who received messages that evoked their own values engaged in deeper processing which ultimately made them favor strong arguments and resist weak ones.

Moral appeals may therefore influence the effectiveness of persuasive communications through multiple channels. They may directly improve the merits of the argument itself (central route), they may serve as identity-based cues and heuristics (peripheral route), or they may increase a person's motivation to scrutinize a message in a more elaborate way (see also Petty and Cacioppo, 1986b). As will be further described below, the present analysis focuses on the influence of moral appeals on argument strength in the context of elaborate processing and the central route to persuasion.

### 4.2.2 Morality and the Potential for Compromise

There are two broad strands of literature in moral psychology that ultimately lead to diverging predictions regarding the effects of moral appeals on argument persuasiveness. Research on *Moral Conviction* conceptualizes moralization as a unique feature of attitude strength (Skitka, Bauman, and Sargis, 2005). According to this view, moral convictions are perceived as “absolutes, or universal standards of truth that others should also share” (Skitka, 2010, 269). As such, moral convictions are viewed by individuals as applying to everyone (universality), they do not require an immediate underlying rationale but are rather seen as facts about the world (objectivity), they can be independent of authority and group norms (autonomy), they elicit strong emotional reactions, and they have an inherent motivational quality (motivation/justification) (Skitka, 2010).

Building on this work, Ryan (2014) argues that moral convictions are not restricted to issues that are traditionally perceived as “moral,” such as abortion or same-sex marriage, but can also include other issues such as economic policies. The degree of moral conviction may therefore vary between individuals as well as across issues. Ryan (2014) further shows that the propensity to moralize—i.e. the tendency to view an issue as a question of “right and wrong”—is related to political participation, extreme political attitudes, arousal of negative emotions, and hostility. In a subsequent study, Ryan (2017) suggests that moralization reorients behavior from maximizing gains to the general adherence to rules. Across multiple studies, the author shows that this tendency translates into stronger opposition to compromise about political issues and decreased support for compromising politicians. These patterns should also translate into attitudes towards—and interactions with—others who hold opposing views. Indeed, moral conviction has been shown to be related to stronger preferences for social distance from (and hostility towards) attitudinally dissimilar others and lower cooperativeness in groups holding heterogeneous views (Skitka, Bauman, and Sargis, 2005). This theoretical perspective therefore ultimately suggests that arguments that emphasize an issue in terms of deeply held moral mandates should entrench people to maintain their prior attitudes and therefore reduce the argument’s persuasiveness.



However, not everyone agrees with this general prediction. In fact, *Moral Foundations Theory* (MFT) offers a more differentiated view regarding the role of moral appeals in facilitating compromise. The theory proposes a taxonomy of basic moral intuitions that is closely related to ideological thinking. According MFT, liberals focus on *individualizing* moral foundations, which include care/harm and fairness/cheating. Conservatives, on the other hand, also emphasize the remaining *binding* foundations of loyalty/betrayal, authority/subversion, and sanctity/degradation (Haidt and Graham, 2007; Graham, Haidt, and Nosek, 2009). Differential emphasis on these moral dimensions is systematically related to attitudes towards a wide variety of divisive political issues (e.g. Koleva et al., 2012; Kertzer et al., 2014; Low and Wui, 2015), personality traits like individual social dominance orientation and right-wing authoritarianism (Federico et al., 2013), as well as voting behavior (Franks and Scherr, 2015). Overall, this body of research suggests that liberals and conservatives endorse different moral foundations and that these differences are closely related to political attitudes, evaluations, and behavior.

An implicit assumption made in this literature is that liberals and conservatives would be more likely to come to agreements *if only they focused on the same moral foundations*. For example Haidt (2012, 365) concludes in his book *The Righteous Mind: Why Good People Are Divided by Politics and Religion*: “Once people join a political team, they get ensnared in its moral matrix. They see confirmation of their grand narrative everywhere, and it’s difficult—perhaps impossible—to convince them that they are wrong *if you argue with them from outside of their matrix*” (emphasis added). In an different article, Graham, Haidt, and Nosek (2009, 1040) contend that their findings “help explain *why liberals and conservatives disagree on so many moral issues* and often find it hard to understand how an ethical person could hold the beliefs of the other side: Liberals and conservatives *base their moral values, judgments, and arguments on different configurations* of the five foundations.”

Several framing studies examining the effects of moral arguments that are congruent with ideological predispositions support this view. For example, binding appeals have been shown to increase recycling behavior among conservatives, whereas individualizing arguments were effective

among liberals (Kidwell, Farmer, and Hardesty, 2013). Similarly, Feinberg and Willer (2013) find that pro-environmental frames emphasizing concerns related to the purity dimension reduce attitudinal gaps of conservatives vis-à-vis liberals. Further studies suggest that morally congruent appeals are effective in shifting attitudes of ideologues on various other issues as well (e.g., Day et al., 2014; Feinberg and Willer, 2015).

Both theories of morality therefore lead to diverging expectations regarding the effect of moral appeals on the potential for compromise: While the moral conviction literature suggests that *any* type of moral appeal should make it harder to overcome disagreements, MFT contends that agreement can be facilitated if two discussants focus on the same underlying moral dimensions. The question whether emphasizing the same foundations can facilitate compromise has important implications—especially in our current political environment. Somewhat surprisingly, however, this claim has not been subjected to a direct empirical test in the context of political discussions.

### 4.2.3 Hypotheses

The structure and dynamics of political discussions can be prohibitively complex, making it difficult to derive clear expectations regarding the persuasiveness of individual arguments and their role in achieving compromise. In order to gain some analytical leverage, consider the following simplified scenario of a conversation between two discussants,  $\mathcal{A}$  and  $\mathcal{B}$ , who disagree on some issue  $x$ . Suppose further that only  $\mathcal{A}$ 's opinion is malleable and may change as an outcome of the discussion.  $\mathcal{B}$ 's own position is firm and she is solely trying to challenge  $\mathcal{A}$ 's view. The conversation begins with  $\mathcal{A}$  making an opening statement describing and defending her opinion—potentially relying on moral justifications.  $\mathcal{B}$  then engages in the discussion and may try to persuade  $\mathcal{A}$  using either moralized or non-moralized arguments. Of course,  $\mathcal{A}$  and  $\mathcal{B}$  can continue to respond to each others' statements until either  $\mathcal{A}$  changes her opinion or the conversation ends without attitude change. Both theoretical perspectives described in the previous section suggest contrasting hypotheses regarding the persuasiveness of  $\mathcal{B}$ 's appeals:

*H1 (Moral Conviction):* Arguments that involve moral appeals will be *less* persuasive than arguments that do not involve moral appeals.

*H2 (Moral Foundations):* Arguments that involve moral appeals will be *more* persuasive than arguments that do not involve moral appeals, but only if they are congruent with the opening statement's moral framework.

To reiterate, in this unidirectional model of a discussion, only  $\mathcal{A}$  stands to maintain or change her prior view, whereas  $\mathcal{B}$  attempts to persuade her discussant. Compromise is achieved in this scenario if  $\mathcal{B}$  is able to persuade  $\mathcal{A}$  to change her attitude. One of the main advantages of this structure is that it enables a clear analytical distinction between statements that are intended as justifications to defend and bolster one's own view (i.e.,  $\mathcal{A}$ 's arguments) and challenges that are targeted to alter a discussant's opinion (i.e.,  $\mathcal{B}$ 's arguments), which is not feasible in a free flowing discussion where—at least potentially—all views are malleable. The following section illustrates how conversations on the Reddit community `/r/ChangeMyView` resemble this stylized conceptualization of a discussion and therefore provide an ideal environment to test both competing hypotheses.

### 4.3 The Subreddit “ChangeMyView”

Reddit is an online discussion board organized into thematic forums called *subreddits*. Users can join these communities based on their interests and each subreddit has its own norms and etiquette that are enforced by voluntary moderators. `/r/ChangeMyView` (CMV) is a subreddit where participants can initiate a discussion by posting an opening statement establishing a personally held view on a particular issue (e.g., “CMV: I believe that the gay marriage discussion isn't as important as the media portrays it to be.”), followed by a brief explanation of their underlying rationale. Other users are then invited to challenge the original poster's (OP) opinion by providing counterarguments. OPs respond to the challenges and—crucially—identify individual posts that changed their mind by awarding a “Delta” ( $\Delta$ ). The community is dedicated to civil discourse—

### 4.3. THE SUBREDDIT “CHANGEMYVIEW”

even for divisive issues—and encourages OPs to be open to changing their views and to award  $\Delta$ s genuinely (see also Jhaver, Vora, and Bruckman, 2017).<sup>2</sup> To date, the subreddit has more than 500,000 subscribed users.

As an illustrative example, consider the following discussion on marriage equality that was posted in 2014. The thread begins with the following opening statement (the posts were slightly edited for readability):

*CMV: I believe that the gay marriage discussion isn't as important as the media portrays it to be.*

The real problem is the concept of marriage itself. In my view, LGBT couples are already married, regardless of the legislation that is imposed on them. Marriage isn't a set of civil rights that confirms your connection to your partner, it's the choice you make to be in private, daily, lifelong commitment to another being.

Tradition dictates that in order to be 'properly' married you have to exchange vows, get a ring, and have a massive celebration (the set of traditions change based upon the culture.) but marriage isn't that, it is simple commitment to another person. The main issue that gay marriage has is that not all couples are given the same civil liberties, but this does not mean that their marriages are void. Marriage isn't decided by bystanders, it's decided by the people who live inside the union. It is for this very reason that a gay couple getting married doesn't affect your own marriage.

I've held this opinion for a while but have never had the opportunity to see if it stood up to criticism. CMV.

Here, the OP argues that marriage equality should be less of a controversy since the defining feature of marriage is the commitment in a relationship rather than its legal status. Several users argued against this view from various perspectives. Below is a sample response that ultimately lead the OP to award a  $\Delta$  to indicate that it changed his or her view:

That would be true if it was just some odd tradition. But it isn't just the ceremony, but also a tax.

Right now there is a gay tax. Gay couples have to pay higher taxes than straight couples because the government gives a tax break for married couples. The reason for this is that married couples tend to be more efficient and better for the government.

<sup>2</sup>The current set of rules for original posts as well as responses can be accessed at <https://www.reddit.com/r/changemyview/wiki/rules>. Additionally, an overview of the current rules is included in Appendix C.1.

The government wants to encourage marriage, so as with all things they encourage they subsidize it.

Gay people provide the exact same benefits to marriage, if not more! Adoption being the largest one.

This tax comes through in multiple ways. The yearly tax and through inheritance. The government doesn't tax inheritance as much for marriage, but if they are simply partners then they get taxed when their "partner" dies.

The state also doesn't allow for gay couples see their loved ones in hospitals or prison because they aren't married.

If this was just in the church I wouldn't care. But this is much more than that.

Note that in principle, the OP could have reacted to this root response by providing additional justifications and the discussion between both users could have continued for a few posts. In this case, the OP directly provided a  $\Delta$ . However, other discussants were less successful in persuading the OP. In contrast to the previous example, the following response did not receive a  $\Delta$ :

If gay marriage is not allowed in a state

1. Their marriages technically *are* null and void, as the state does not recognize them.
2. Marriage is not actually decided by the people in the union, since there are legal requirements as well as legal benefits. Which brings me to my next point.
3. There are several legal benefits (as well as tax benefits) to being married. States which do not allow gay marriage do not give these legal benefits to gay couples.

You might believe you are married to someone, but the term "marriage" is a political one indeed since it has legal ramifications.

While both responses emphasize the importance of legal considerations in justifying the need for marriage equality, only one of the contributions persuaded the OP sufficiently such that she awarded a  $\Delta$ .

This online format provides an ideal opportunity to explore the correlates of argument persuasiveness consistent with the stylized structure outlined in the previous section. Discussions begin with a short explanation of a person's opinion on a given topic. Multiple users attempt to

counterargue the OP's point of view from various perspectives in a civil dialogue. Most importantly, OPs explicitly identify and label arguments they deemed persuasive enough to change their views. The nature of the conversations on CMV as well as the anonymity of individual users turns the focus on the content and merits of arguments (i.e., the central route to persuasion) rather than source cues and identity-related factors. In contrast to past framing studies which usually implement single messages, users on CMV evaluate a multitude of available arguments, which allows for a unique counterfactual design to study persuasive messages that can be directly linked to the OP's initial justifications. Lastly, examining discussions on CMV allows for an examination of a wide array of issues.

Recent research in machine learning and computational linguistics has started to use CMV to study online discussions (Wei, Liu, and Li, 2016; Hidey et al., 2017). The following analyses leverage a set of matched argument pairs extracted from CMV by Tan et al. (2016), who explore interaction dynamics on CMV by analyzing linguistic features (such as, for example, the use of personal pronouns) that predict persuasiveness as well as the malleability of original posts. Their dataset includes more than 10,000 discussions that were posted on the subreddit between January 2013 and May 2015. It is important to note that the analysis published by Tan et al. (2016) focuses less on the content of discussions (i.e., *what is being said*) but rather examined discussion dynamics and linguistic characteristics (i.e., *how it is expressed*) to predict persuasiveness. The following analyses explicitly turn to the effects of moral content on discussion outcomes.

## 4.4 Opinion Change in Online Discussions

Consider again the simplified model of a discussion between person  $\mathcal{A}$  and  $\mathcal{B}$ , where  $\mathcal{A}$  stands to defend her view against the challenges put forward by  $\mathcal{B}$ . While the hypotheses specified above are focused on the persuasiveness of  $\mathcal{B}$ 's arguments (i.e., discussion posts that respond to the OP in the context of CMV), it is helpful to first explore the opening statements initiating each discussion and examine the extent to which OPs are willing to award  $\Delta$ s in the first place.

To provide an initial overview of the range of topics that are covered in the set of 10,000 initial statements included in the data, I extract 20 clusters of co-occurring terms via Latent Dirichlet Allocation (Blei, Ng, and Jordan, 2003). The topic model is solely based on contents of the original posts starting each discussion thread (disregarding subsequent comments). Statements were pre-processed by removing numbers, punctuation, symbols, hyphens, URLs, as well as stopwords. All remaining terms were stemmed and only included if they appeared in at least 10 different posts. Figure 4.1 presents the average topic proportions across opening statements based on the model. For each topic, the plot additionally displays the ten most likely word stems as well as a descriptive label on the y-axis.

Conversations on CMV range across a variety of topics such as economic issues, gender/sexuality, or domestic and international politics. Of course, it could be argued that some of these topics—for example those related to religion—more easily lend themselves to concerns about morality. Notwithstanding, recent work in moral psychology by Ryan (2014) and others routinely emphasizes that in principle, any issue bears the potential to be moralized by individuals. However, in an effort to preclude any concerns about potential confounding effects related to topic selection, the main analyses reported below focuses on comparing arguments *within* a given discussion thread.

The Internet is not necessarily known as a place where people are willing to change their mind about *any* issue. Yet, CMV maintains an open atmosphere that encourages users to acknowledge arguments that change their perspective. The rules of the subreddit state that users should “Award a delta if you’ve acknowledged a change in your view. [...] Please note that a delta is not a sign of ‘defeat’, it is just a token of appreciation towards a user who helped tweak or reshape your opinion. A delta also doesn’t mean the discussion has ended.”<sup>3</sup> Of course, this does not imply that every OP awards a  $\Delta$  throughout a conversation. Figure 4.2 displays the number of discussion threads included in the dataset where OPs indicated that one of the responses changed their mind.

In about two thirds of discussions on CMV between 2013 and 2015, OPs did not award a

---

<sup>3</sup> <https://www.reddit.com/r/changemyview/wiki/rules>, last accessed April 22, 2018

#### 4.4. OPINION CHANGE IN ONLINE DISCUSSIONS

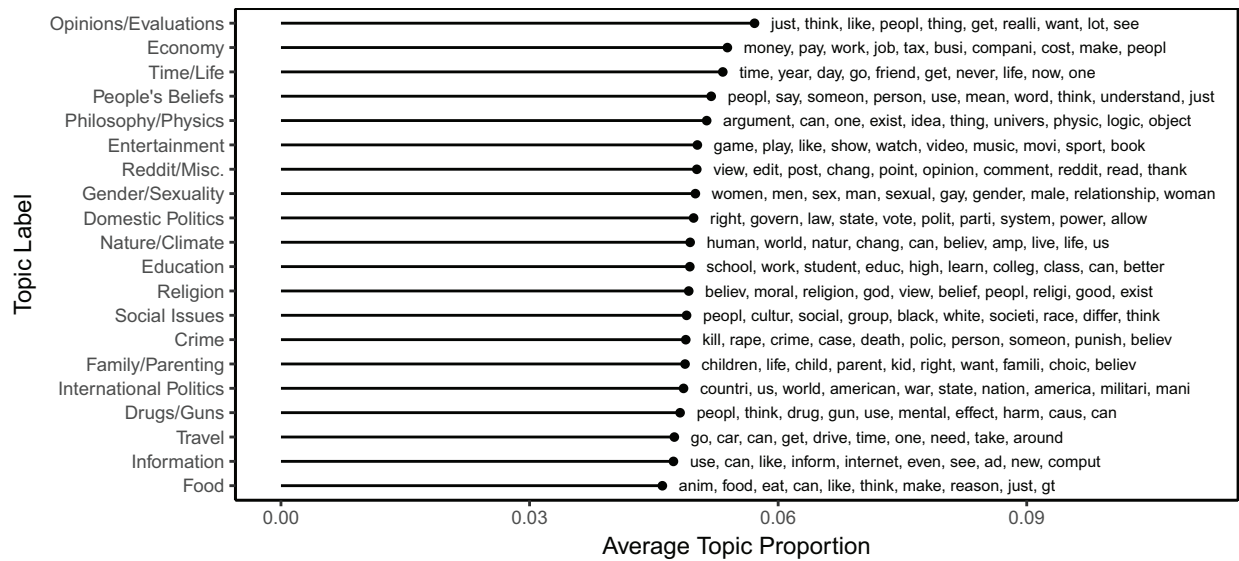


Figure 4.1: Average topic proportions in opening statements on /r/ChangeMyView/ based on a basic LDA model with 20 topics. The plot additionally displays the ten most likely terms associated with each respective topic.

△ for any of the counterarguments that were put forward, which leaves about 3,000 individual threads where OPs indicated that at least one of the responses changed their views. Interestingly, additional analyses included in the supplementary material show that there are only minimal differences in topic proportions between discussion threads that resulted in persuasion versus threads that did not (see Appendix C.4).

In their original study, Tan et al. (2016) mainly investigated linguistic patterns (e.g., use of personal pronouns) and differences in style (formatting etc.) that predicted resistance to persuasion among OPs. They conclude for instance that “comparative adjectives and adverbs are a sign of malleability, while superlative adjectives suggest stubbornness.” The goal of the present analysis, in turn, is to go beyond linguistic patterns that are unrelated to content and explore the role of moral appeals in facilitating or inhibiting compromise. In order to capture moralized arguments, I rely on the MFT dictionary proposed by Graham, Haidt, and Nosek (2009), which contains lists of word stems that signal each of the five moral foundations (care, fairness, loyalty, authority, sanctity) as well as a category of general moral terms.<sup>4</sup>

<sup>4</sup>See Appendix C.2 for the complete dictionary.



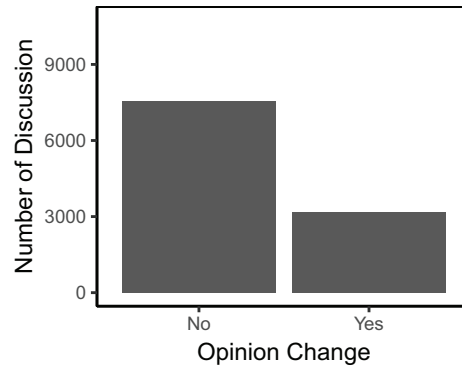


Figure 4.2: Number of discussions on /r/ChangeMyView/ that resulted in opinion change (at least one  $\Delta$  awarded by OP) versus not.

Figure 4.3 displays the percentage of dictionary terms for each foundation in the opening statements initiating a discussion on CMV (in proportion to the total number of words in each post). The plot compares the reliance on moral terms between OPs that subsequently changed their view versus OPs that did not. As an initial observation, it is interesting to note that the distribution of dictionary terms across foundations is strikingly similar to the proportions of moral terms in open-ended responses to the likes-dislikes questions included in the American National Election study (c.f., Kraft, 2018): The most prevalent dimensions are *care* and *authority*, while occurrences of *sanctity* are largely negligible. Observing these similarities is noteworthy since they are suggestive of a common mechanism driving the emphasis on moral considerations when justifying preferences in a public opinion survey as well as in online discussions.

More important for the purposes of this paper is the fact that the percentage of dictionary terms across foundations appears smaller among opening statements that resulted in opinion change than among those that did not. More specifically, OPs who did not award any  $\Delta$ s in the subsequent discussion put a significantly stronger emphasis on moral considerations related to loyalty and authority ( $p < .01$  in both cases after accounting for multiple comparisons using Bonferroni correction). Similar results can be obtained after aggregating all moral dictionary terms in a single category: OPs who were not persuadable on CMV use more words related to morality overall than OPs who indicate that the discussion changed their view ( $p < .001$ ).

#### 4.4. OPINION CHANGE IN ONLINE DISCUSSIONS

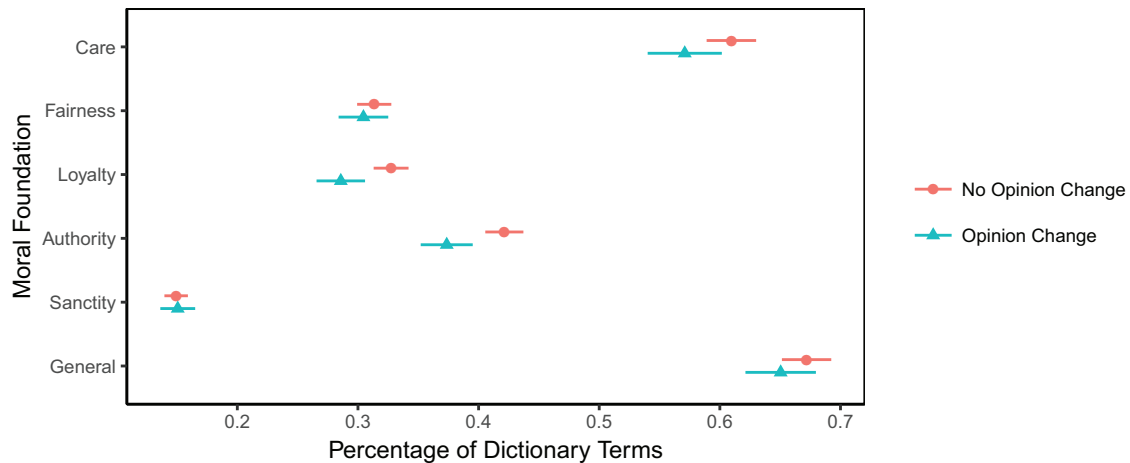


Figure 4.3: Moral foundations and persuadability: Average percentage of dictionary terms relative to the total number of words in each opening statement beginning a discussion, comparing discussions where the OP subsequently awarded a  $\Delta$  (opinion change) or not (including 95% confidence intervals).

At first look, the findings appear consistent with the moral conviction literature, which posits that people who hold moralized attitudes are less willing to compromise and deviate from their prior beliefs (e.g., Skitka, Bauman, and Sargis, 2005; Ryan, 2014, 2017). Yet, there are important issues that make it difficult to draw strong conclusions based on these initial results. First of all, there may be unobserved confounding factors that are related to both the OPs willingness to award  $\Delta$ s as well as the chosen discussion topic (which could be more or less aligned with moral considerations). The content of opening statements may also induce selection bias in user responses which can impact the nature of their comments and ultimately the productivity of discussions. Furthermore, there is no way of contrasting the potentially diverging impact of morally congruent arguments by exclusively examining the malleability of initial opinions. The following analyses address these problems by comparing the relative persuasiveness of arguments *within a given discussion thread*.

## 4.5 What Makes an Argument Persuasive?

The previous section demonstrated that the OPs' reliance on moral language in opening statements is inversely related to their willingness to change their view in the subsequent discussion. Now I focus directly on the persuasiveness of comments that are made in response to a given opening statement on CMV. In the context of the simplified discussion framework outlined above, I examine the arguments brought forward by  $\mathcal{B}$ , who is challenging  $\mathcal{A}$ 's view. This allows me to directly compare the moral conviction hypothesis with the moral foundations hypothesis, which have diverging predictions regarding the effectiveness of moralized appeals in discussions. Note that the arguments presented by  $\mathcal{B}$  do not only include her initial post (i.e., the root response), but also any subsequent posts that are mentioned in the evolving discussion between  $\mathcal{A}$  and  $\mathcal{B}$  (i.e., the full response path).

In the original analysis by Tan et al. (2016), the authors implement a simple method to select pairs of arguments that respond to the same original post, with only one of the selected responses being successful in changing the OPs view. While differing in persuasiveness, arguments are matched in such a way that they are as similar as possible in terms of their word choice. More specifically, Tan et al. (2016) select argument pairs by maximizing their Jaccard similarity:

$$\text{Jaccard}(B_{\Delta}, B_{-\Delta}) = \frac{|B_{\Delta} \cap B_{-\Delta}|}{|B_{\Delta} \cup B_{-\Delta}|}, \quad (4.1)$$

where  $B_{\Delta}$  and  $B_{-\Delta}$  are the sets of words in two response paths associated with the same opening statement (one receiving a  $\Delta$ , the other not). In other words, they match each successful counterargument to an unsuccessful response that shares the largest proportion of common words (disregarding stopwords). As Tan et al. (2016, 617) describe: "This leads to a balanced binary prediction task: which of the two lexically similar rooted path-units is the successful one?"<sup>5</sup>

<sup>5</sup>As additional selection criteria and to avoid trivial posts, arguments are removed if they are shorter than 50 words, only include clarifying questions, or if the opening statement received fewer than 10 responses overall and fewer than 3 unsuccessful challenges (see Tan et al., 2016, 617 for details).

## 4.5. WHAT MAKES AN ARGUMENT PERSUASIVE?

The analyses reported below rely on this approach to select matched argument pairs for comparison. To reiterate, I focus on discussions in which OPs awarded at least one  $\Delta$ . A response that received a  $\Delta$  is then matched to another argument within the same discussion that was not successful in changing the OP's view but is as similar as possible in terms of its vocabulary. Note that in principle, this strategy should make it more difficult to find differences in the MFT dictionaries as argument pairs are matched based on their lexical similarity. One might worry, however, that the necessary initial selection on discussions where OPs ultimately awarded at least one  $\Delta$  might disproportionately discard cases where the initial statement emphasized moral considerations. Luckily, that is not the case. Figure 4.4 shows that almost all of the opening statements in the matched pair selection mention at least one of the moral dictionary terms. Furthermore, the proportion of moral dictionary terms among this set of opening statement shows the same pattern as in Figure 4.3 (results included in Appendix C.3).

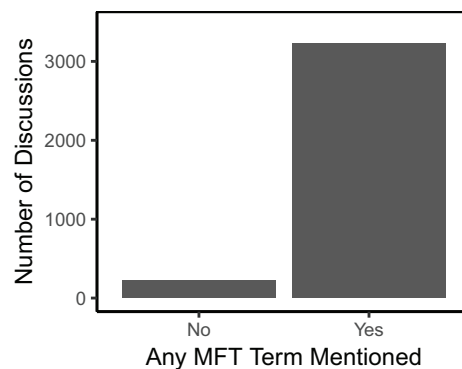


Figure 4.4: Number of opening statements in the paired argument data that included *any* term mentioned in the MFT dictionary.

An important unresolved issue using this approach is that the matching procedure only focuses on the set of unique words that are used in each response path and does not take into account their relative length. This can be especially problematic since persuasive discussions tend to be longer and involve at least a few back-and-forth exchanges between the OP and the challenger (c.f., Tan et al., 2016, 616). Figure 4.5 displays the distribution of the differences in word counts between successful and unsuccessful argument pairs. Clearly, longer responses are more likely to

be awarded a  $\Delta$ , which might jeopardize potential inferences about the relative reliance on moral dictionary terms.

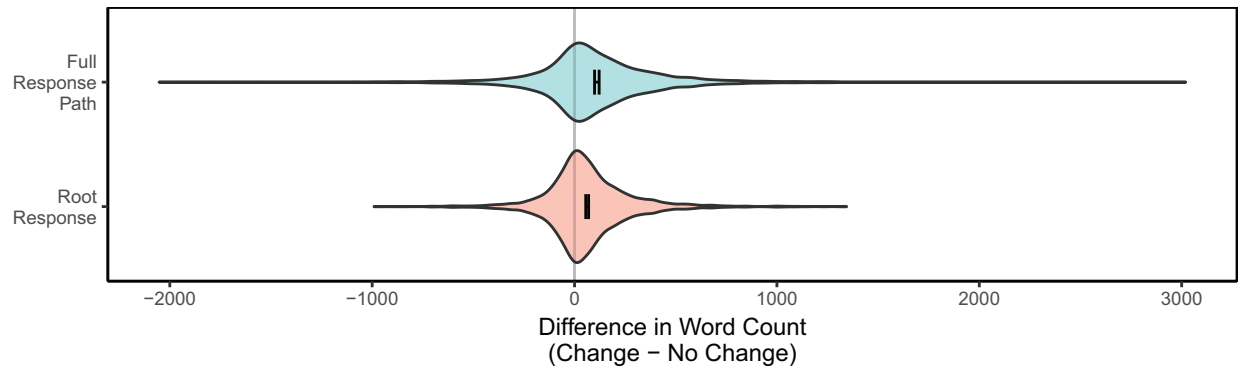


Figure 4.5: Difference in response lengths between successful and unsuccessful counterarguments. The narrow black bars display the 95% confidence interval of mean differences.

As a first step to alleviate this concern, it is worth noting that I only examine the *percentages* of MFT dictionary occurrences in each discussion rather than a raw count, which implies that the prevalence of moral considerations is standardized by the overall length of each post. Notwithstanding, I take additional precautions proposed by Tan et al. (2016) to check the robustness of the results. First, I not only examine differences when looking at the entire response path of a discussion between two users (i.e., all posts that were part of the dialogue with the OP), but restrict the analysis to focus on the challenger’s root response to the opening statement only. As can be seen in Figure 4.5, the differences in word counts between argument pairs are significantly smaller. Recovering the same patterns in the root response as in the full response path indicates that the initial arguments that triggered an exchange with the OP are by themselves predictive of the outcome of the discussion. To be fair, there are still substantial differences in the length of successful versus unsuccessful root responses. As a second robustness check, I additionally truncate the longer root response of a given pair as follows: I count the total number of words in each post and simply cut off the end of the longer response such that both word counts in a given pair are exactly equal. While this approach eliminates any concerns about argument length as a confounding factor, it comes at the price of losing a lot of information by ignoring potentially

valuable content. Using this framework, I now turn to the analysis of the persuasiveness of moral arguments made in discussions on CMV.

### 4.5.1 Moral Appeals are Futile...

Recall that the moral conviction hypothesis posits that moralized arguments will be less persuasive than arguments that do not involve moral appeals. In order to test this proposition, I examine the argument pairs matched within discussions and compare MFT dictionary proportions between contributions that were successful in receiving a  $\Delta$  and those that were unsuccessful. Figure 4.6 presents the results of logistic regressions predicting an argument's success in triggering opinion change as a function of moral language use.

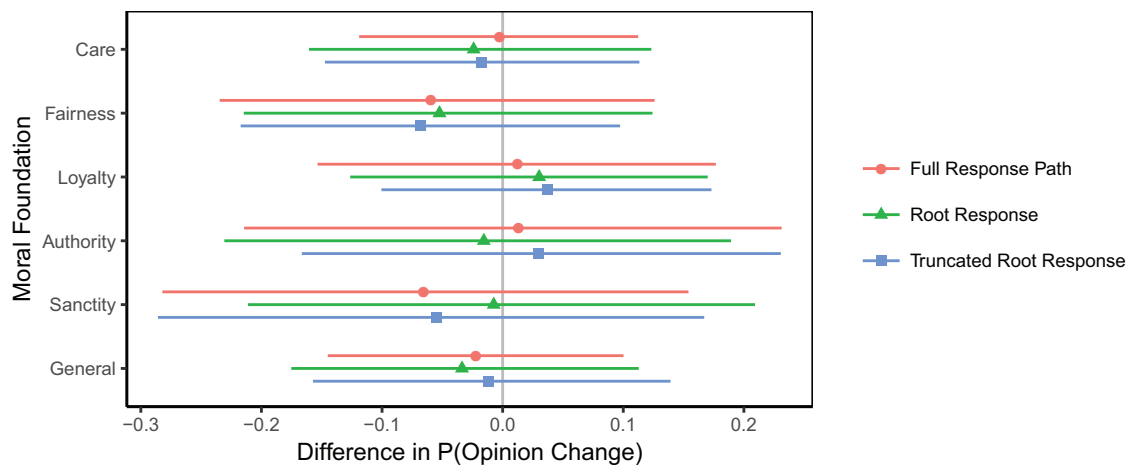


Figure 4.6: Moral foundations and persuasiveness: Change in predicted probability of opinion change ( $\Delta$  awarded) when MFT dictionary term proportions are increased from their minimum (no moral terms mentioned) to their empirical maximum, holding the remaining foundations constant at their mean (including 95% confidence intervals). Estimates are based on logit models with standard errors clustered by discussion thread. Full model results are displayed in the appendix, Table C.1.

The figure displays the marginal effect on the probability of opinion change when increasing MFT dictionary proportions for each foundation from zero to their respective empirical maximum. Positive values indicate that arguments with larger proportions of dictionary terms belonging to a given foundation have a higher probability of receiving a  $\Delta$ . Again, according to the literature

on moral conviction, we would expect the opposite, namely that arguments focusing on moral considerations should be less persuasive. As discussed in the previous section, the analyses are implemented for the full response path as well as focusing only on (truncated) root responses.

The results show that evoking moral considerations in counterarguments does not affect the likelihood of changing the OPs' view on a given issue. This finding furthermore holds after combining all dictionary term proportions in an aggregate measure of moralization across foundations ( $p > .45$ ). Moralized arguments as such are therefore no less persuasive and do not reduce compromise, a result that is not consistent with the moral conviction literature.

### 4.5.2 ...Unless We're Speaking the Same Moral Language

In contrast to the moral conviction hypothesis, moral foundations theory suggests that we cannot fully understand the effect of moral appeals without taking into account the discussion partner's moral framework. What is decisive from this perspective is the congruence in moral arguments between both discussants. I measure the moral congruence between an OP's opening statement and each counterargument by computing the cosine similarity in their respective MFT dictionary scores. In general, using cosine similarities based on vectors of word counts is a standard approach in text analysis to quantify the similarity of documents independent of their length (e.g., Manning et al., 2008). More formally, moral congruence can therefore be written as:

$$\text{MFT Congruence} = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}, \quad (4.2)$$

where  $\vec{a}$  is the vector of dictionary counts in the OP's opening statement and  $\vec{b}$  is the respective vector for a response. The measure ranges from 0 (no moral overlap) to 1 (equal emphasis on the same moral foundations). Moral congruence is also set to zero if either one of the statements does not contain a single term included in the dictionary.

To reiterate, the moral foundations hypothesis posits that arguments involving moral appeals

#### 4.5. WHAT MAKES AN ARGUMENT PERSUASIVE?

will be more persuasive than arguments that do not involve moral appeals, but only if they are congruent with the opening statement's moral framework. In contrast, the moral conviction literature would predict a negative effect of moral congruence, since it implies that both discussants, who hold opposing views on an issue, use moralized arguments that ultimately reduce the potential for compromise. Figure 4.7 displays the effect of moral congruence on argument success. Similar to the previous analysis presented in Figure 4.6, it shows the predicted change in the probability of persuasion based on a logistic regression model, but now examining the effect of an increase in MFT congruence from its minimum to its maximum. Positive values indicate that posts were more likely to be awarded a  $\Delta$  by the OP if they used language that is morally congruent with the OP's opening statement. On average, emphasizing the same moral foundations as the opening statement (as compared to no overlap in moral language at all) increases the probability of opinion change by about 7 percentage points. These results are consistent with MFT as moral congruence is associated with a higher probability of opinion change.

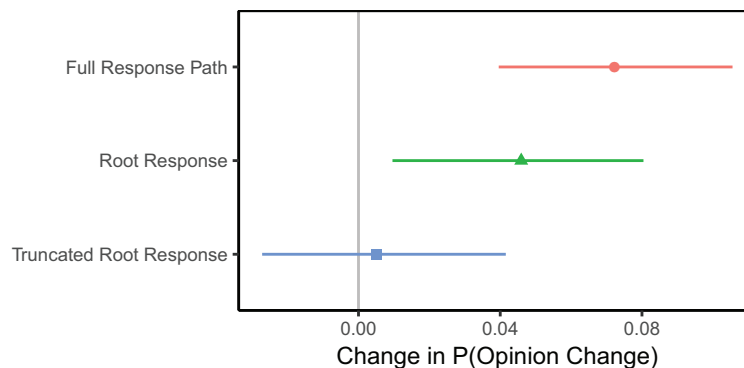


Figure 4.7: Moral congruence and persuasiveness: Change in predicted probability of opinion change ( $\Delta$  awarded) when MFT congruence is increased from its minimum (no overlap in moral terms) to their empirical maximum (equal emphasis on each moral foundation) (including 95% confidence intervals). Estimates are based on logit models with standard errors clustered by discussion thread. Full model results are displayed in the appendix, Table C.2.

The positive relationship between moral congruence and persuasiveness remains significant irrespective of whether I examine the content of the entire discussion (full response path), or restrict the analysis to each user's first post challenging the OP (root response). However, the



mean difference in moral congruence does not reach conventional levels of statistical significance after truncating root responses to the same length for each pair. This finding could either suggest that the truncation procedure introduces too much noise to recover any differences, or alternatively that the measure of moral congruence is confounded by the differential length of successful and unsuccessful posts. The former seems less likely to be an issue because I recovered only marginal differences in raw dictionary term percentages between raw and truncated rooted responses in the previous section.

It is also important to emphasize that the higher moral congruence among persuasive posts is by no means driven by the fact that successful arguments use similar language to the opening statement *in general*. Quite contrary, Tan et al. (2016, 618) concluded in their study that when looking at the entire vocabulary of responses (excluding stopwords), then persuasive arguments used significantly more *different wording* than original post. In other words, a similar general vocabulary across all words is less persuasive, whereas a similar use of terms belonging to each moral foundation proved to be more persuasive. As such, the results presented here appear to capture the unique persuasive effect of morally congruent arguments.

### 4.6 Conclusion

Political elites on both sides of the aisle routinely rely on moral rhetoric in order to bolster their views, which induces strong emotional reactions among citizens (Lipsitz, 2018) and can ultimately influence their attitudes (e.g., Clifford and Jerit, 2013; Clifford et al., 2015). As such, it does not seem surprising that the increasingly partisan and polarized environment in the United States has been linked to stronger tendencies among citizens to moralize politics (Garrett and Bankert, 2018). Is the only solution to overcome this trend to de-emphasize moral convictions when discussing political issues? Or is it rather the case that morality may even be helpful in overcoming disagreements as long as people rely on the same moral frameworks?

The present paper addresses these questions by contrasting two strands of research in moral

psychology that lead to diverging predictions regarding the role of morality in political compromise. Previous work on moral conviction suggests that individuals who moralize politics should be less willing to compromise and therefore resist persuasion through moral appeals. On the other hand, moral foundations theory posits that compromise is indeed possible as long as the discussants use the same moral language.

Both competing hypotheses are tested by relying on a unique dataset of online discussions on the Reddit community CMV compiled by Tan et al. (2016). Overall, the empirical patterns support moral foundations theory and stand in contrast to predictions rooted in the literature on moral conviction. While general levels of moralization have little impact on argument persuasiveness, the results show that an argument's moral congruence with the discussant's opening statement increases the likelihood of changing his or her view. As such, moral appeals can facilitate compromise and change people's minds as long as they are consistent with their existing moral frameworks. Rather than automatically driving people further apart, moral appeals might therefore help bridge the growing divide between liberals and conservatives. More broadly, the paper shows that the field of moral psychology stands to benefit from a further integration of two prominent theoretical frameworks that developed largely independent of each other and—unfortunately—still exhibit relatively little interconnections.

At the same time, the analysis presented here has important limitations. One of the biggest potential issues is the fact that the matched argument pairs differ in length, which may confound the relationship between morality and persuasiveness. I addressed this concern by only examining measures that are standardized by the total number of words in each post and by examining root responses in addition to full response paths. The results are largely robust to these varying specifications, with the important exception of the effect of moral congruence in the truncated root response. More generally, while it is a substantial advantage that the discussions on CMV cover a wide range of topics, it can be argued that some of them are ultimately irrelevant for moral considerations (such as software and technology).<sup>6</sup> On the other hand, such inherently non-

<sup>6</sup>For example, users discussed the merits of the Windows operating system as compared to macOS X, which

## CHAPTER 4. CHANGE MY VIEW

---

political and non-moral discussions should not induce any systematic biases between successful and non-successful arguments. For the purpose of this paper, I decided against filtering out subsets of discussions based on their thematic relevance since such a strategy would raise concerns about potential selection bias. I leave it to future research to leverage more controlled environments and focus on specific (political) issues—for example in the context of laboratory experiments. In contrast to framing studies conducted in the past, however, it is time to open the black box of conversations and directly examine the content of discussions in order to better understand the mechanisms underlying attitude change, persuasion, and compromise.

---

is obviously fruitless since Linux is superior to either of them and anyone who says otherwise is morally wrong.

## Concluding Remarks

What can we learn from the way citizens describe and rationalize their political beliefs in surveys and social interactions? This dissertation provides three examples that showcase how research in political psychology and public opinion can leverage open-ended and text-based measures in order to provide new avenues to test important hypotheses of political cognition, attitude formation, and social influence. Thus, my research contributes to the literature both methodologically as well as theoretically.

This first empirical chapter proposes a simple but powerful framework to measure political sophistication based on the complexity of individual attitude expression in open-ended survey responses. Discursive sophistication is shown to be a better predictor of important competences in the realm of politics than conventional political knowledge metrics. The chapter then illustrates how the measure helps refine previous insights about the oft-cited gender gap in political knowledge. Women might know fewer facts about institutions and elites, but they do not differ substantively in the sophistication of their expressed political beliefs.

The second empirical chapter proceeds by examining the content of open-ended survey responses to explore whether and how individuals evoke moral considerations when discussing their political beliefs. Using a previously validated dictionary, I find systematic ideological differences in moral reasoning—even when respondents are not explicitly asked about morality. The results

## CHAPTER 5. CONCLUDING REMARKS

---

further suggest that the reliance on moral considerations in attitude expression is amplified by the moral content of individual media environments.

The last empirical chapter explores potential implications of moralized attitude expression in the context of discussions. Drawing on a unique dataset from the online discussion board *Reddit*, this part of the dissertation examines how moral appeals can affect argument persuasiveness and the likelihood of compromise. The findings indicate that moral arguments can facilitate compromise, but only to the extent that they are congruent with the moral framework of the opposing discussant.

In summary, the results discussed throughout this dissertation have broader implications for the study of public opinion. Contemporary American politics is frequently characterized by increasing conflict and polarization among partisans. A better understanding of the nature of political attitude expression and especially the conditions under which discussions can lead to compromise improves our understanding of the social antecedents of polarization. In order to address these and other important issues facing political science as a discipline and society as a whole, it is time for researchers to listen more closely to what people have to say.

# Chapter 6

## Bibliography

- Abramowitz, Alan I, and Kyle L Saunders. 2008. "Is polarization a myth?" *The Journal of Politics* 70 (02): 542–555.
- Ahn, Toh-Kyeong, Robert Huckfeldt, and John Barry Ryan. 2010. "Communication, influence, and informational asymmetries among voters." *Political Psychology* 31 (5): 763–787.
- Barabas, Jason. 2004. "How deliberation affects policy opinions." *American Political Science Review* 98 (04): 687–701.
- Barabas, Jason, Jennifer Jerit, William Pollock, and Carlisle Rainey. 2014. "The Question(s) of Political Knowledge." *American Political Science Review* 108 (04): 840–855.
- Bartels, Larry M. 2005. "Homer gets a tax cut: Inequality and public policy in the American mind." *Perspectives on Politics* 3 (1): 15–31.
- Basinger, Scott J, and Howard Lavine. 2005. "Ambivalence, information, and electoral choice." *American Political Science Review* 99 (02): 169–184.
- Beck, Paul Allen, Russell J Dalton, Steven Greene, and Robert Huckfeldt. 2002. "The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices." *American Political Science Review* 96 (01): 57–73.
- Benoit, Kenneth, Kevin Munger, and Arthur Spirling. 2017. "Measuring and Explaining Political Sophistication Through Textual Complexity." Available at SSRN: <https://ssrn.com/abstract=3062061> or <http://dx.doi.org/10.2139/ssrn.3062061>.
- Benoit, Kenneth, Michael Laver, and Slava Mikhaylov. 2009. "Treating words as data with error: Uncertainty in text statements of policy positions." *American Journal of Political Science* 53 (2): 495–513.
- Blei, David M, Andrew Y Ng, and Michael I Jordan. 2003. "Latent dirichlet allocation." *the Journal of machine Learning research* 3: 993–1022.
- Bolsen, Toby, James N. Druckman, and Fay Lomax Cook. 2014a. "The Influence of Partisan Motivated Reasoning on Public Opinion." *Political Behavior* 36 (2): 235–262.

## CHAPTER 6. BIBLIOGRAPHY

---

- Cappella, Joseph N, Vincent Price, and Lilach Nir. 2002. "Argument repertoire as a reliable and valid measure of opinion quality: Electronic dialogue during campaign 2000." *Political Communication* 19 (1): 73–93.
- Chaiken, S., and A. H. Eagly. 1989. "Heuristic and Systematic Information Processing Within and Beyond the Persuasion Context." In *Unintended Thought*, ed. J. S. Uleman and J. A. Bargh. Vol. 212. New York: Guilford Press.
- Chong, Dennis. 1993. "How people think, reason, and feel about rights and liberties." *American Journal of Political Science* 37 (3): 867–899.
- Chong, Dennis, and James N. Druckman. 2007. "Framing theory." *Annual Review of Political Science* 10: 103–126.
- Clifford, Scott. 2014. "Linking Issue Stances and Trait Inferences: A Theory of Moral Exemplification." *Journal of Politics* 76 (3): 698–710.
- Clifford, Scott, and Jennifer Jerit. 2013. "How Words Do the Work of Politics: Moral Foundations Theory and the Debate over Stem Cell Research." *Journal of Politics* 75 (3): 659–671.
- Clifford, Scott, and Jennifer Jerit. 2016. "Cheating on Political Knowledge Questions in Online Surveys: An Assessment of the Problem and Solutions." *Public Opinion Quarterly*.
- Clifford, Scott, and Jennifer Jerit. 2018. "Disgust, Anxiety, and Political Learning in the Face of Threat." *American Journal of Political Science* (forthcoming): <https://doi.org/10.1111/ajps.12350>.
- Clifford, Scott, Jennifer Jerit, Carlisle Rainey, and Matt Motyl. 2015. "Moral Concerns and Policy Attitudes: Investigating the Influence of Elite Rhetoric." *Political Communication* 32 (2): 229–248.
- Cobb, Michael D, and James H Kuklinski. 1997. "Changing minds: Political arguments and political persuasion." *American Journal of Political Science* 41 (1): 88–121.
- Colombo, Céline. 2016. "Justifications and Citizen Competence in Direct Democracy: A Multi-level Analysis." *British Journal of Political Science*: 1–20.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: Free Press.
- Cramer, Katherine J, and Benjamin Toff. 2017. "The Fact of Experience: Rethinking Political Knowledge and Civic Competence." *Perspectives on Politics* 15 (3): 754–770.
- 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): 1559–1573.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, London: Yale University Press.

- 
- Denny, Matthew J., and Arthur Spirling. 2018. "Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It." *Political Analysis* 26 (2): 168–189.
- Diermeier, Daniel, Jean-François Godbout, Bei Yu, and Stefan Kaufmann. 2012. "Language and ideology in Congress." *British Journal of Political Science* 42 (01): 31–55.
- Dolan, Kathleen. 2011. "Do women and men know different things? Measuring gender differences in political knowledge." *The Journal of Politics* 73 (01): 97–107.
- Dow, Jay K. 2009. "Gender differences in political knowledge: Distinguishing characteristics-based and returns-based differences." *Political Behavior* 31 (1): 117–136.
- Druckman, James N. 2014. "Pathologies of studying public opinion, political communication, and democratic responsiveness." *Political Communication* 31 (3): 467–492.
- Druckman, James N, and Kjersten R Nelson. 2003. "Framing and deliberation: How citizens' conversations limit elite influence." *American Journal of Political Science* 47 (4): 729–745.
- Eggers, Andrew C, and Arthur Spirling. 2014. "Ministerial Responsiveness in Westminster Systems: Institutional Choices and House of Commons Debate, 1832–1915." *American Journal of Political Science* 58 (4): 873–887.
- Eggers, Andrew C., and Arthur Spirling. 2018. "The Shadow Cabinet in Westminster Systems. Modeling Opposition Agenda Setting in the House of Commons, 1832–1915." *British Journal of Political Science* 48 (2): 343–367.
- Federico, Christopher M., Christopher R. Weber, Damla Ergun, and Corrie Hunt. 2013. "Mapping the Connections between Politics and Morality: The Multiple Sociopolitical Orientations Involved in Moral Intuition." *Political Psychology* 34 (4): 589–610.
- Feinberg, Matthew, and Robb Willer. 2013. "The moral roots of environmental attitudes." *Psychological Science* 24 (1): 56–62.
- Feinberg, Matthew, and Robb Willer. 2015. "From Gulf to Bridge When Do Moral Arguments Facilitate Political Influence?" *Personality and Social Psychology Bulletin* 41 (12): 1665–1681.
- Fortin-Rittberger, Jessica. 2016. "Cross-National Gender Gaps in Political Knowledge How Much Is Due to Context?" *Political Research Quarterly* 69 (3): 391–402.
- Fraile, Marta. 2014. "Does deliberation contribute to decreasing the gender gap in knowledge?" *European Union Politics* 15 (3): 372–388.
- Franks, Andrew S., and Kyle C. Scherr. 2015. "Using Moral Foundations to Predict Voting Behavior: Regression Models from the 2012 US Presidential Election." *Analyses of Social Issues and Public Policy* 15 (1): 213–231.
- Garrett, Kristin N, and Alexa Bankert. 2018. "The Moral Roots of Partisan Division: How Moral Conviction Heightens Affective Polarization." *British Journal of Political Science*: 1–20.



## CHAPTER 6. BIBLIOGRAPHY

---

- Geer, John G. 1988. "What do open-ended questions measure?" *Public Opinion Quarterly* 52 (3): 365–367.
- Geer, John G. 1991. "Do open-ended questions measure salient issues?" *Public Opinion Quarterly* 55 (3): 360–370.
- Gilens, Martin. 2001. "Political ignorance and collective policy preferences." *American Political Science Review* 95 (02): 379–396.
- Graber, Doris A. 2001. *Processing politics: Learning from television in the Internet age*. University of Chicago Press.
- Graham, Jesse, Brian A. Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter H. Ditto. 2011. "Mapping the Moral Domain." *Journal of Personality and Social Psychology* 101 (2): 366–385.
- Graham, Jesse, Jonathan Haidt, and Brian A. Nosek. 2009. "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology* 96 (5): 1029–1046.
- Graham, Jesse, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, S. Wojcik, and Peter H. Ditto. 2013. "Moral Foundations Theory: The pragmatic Validity of Moral Pluralism." *Advances in Experimental Social Psychology* 47: 55–130.
- Grice, Herbert Paul. 1975. "Logic and conversation." In *Syntax and semantics 3: Speech arts*, ed. Peter Cole, Jerry L Morgan, and John P Kimball. New York: Academic Press.
- Grice, Herbert Paul. 1978. "Further notes on logic and conversation." In *Syntax and semantics 9: Pragmatics*, ed. Peter Cole. New York: Academic Press.
- Grimmer, Justin. 2010. "A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases." *Political Analysis* 18 (1): 1–35.
- Haidt, Jonathan. 2007. "The new synthesis in moral psychology." *science* 316 (5827): 998–1002.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Random House.
- Haidt, Jonathan, and Jesse Graham. 2007. "When Morality Opposes Justice: Conservatives Have Moral Intuitions that Liberals May Not Recognize." *Social Justice Research* 20 (1): 98–116.
- Hetherington, Marc J. 2001. "Resurgent mass partisanship: The role of elite polarization." *American Political Science Review* 95 (03): 619–631.
- Hidey, Christopher, Elena Musi, Alyssa Hwang, Smaranda Muresan, and Kathy McKeown. 2017. "Analyzing the Semantic Types of Claims and Premises in an Online Persuasive Forum." In *Proceedings of the 4th Workshop on Argument Mining*. pp. 11–21.

- 
- Huckfeldt, Robert, Paul Allen Beck, Russell J Dalton, and Jeffrey Levine. 1995. "Political environments, cohesive social groups, and the communication of public opinion." *American Journal of Political Science* 39: 1025–1054.
- Huddy, Leonie, Lilliana Mason, and Lene Aarøe. 2015. "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity." *American Political Science Review* 109 (01): 1–17.
- Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, not ideology. A social identity perspective on polarization." *Public Opinion Quarterly* 76 (3): 405–431.
- Iyengar, Shanto, and Sean J Westwood. 2015. "Fear and Loathing Across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science* 59 (3): 690–707.
- Iyer, Ravi, Jesse Graham, Spassena Koleva, Peter Ditto, and Jonathan Haidt. 2010. "Beyond Identity Politics: Moral Psychology and the 2008 Democratic Primary." *Analyses of Social Issues and Public Policy* 10 (1): 293–306.
- Jacoby, William G. 2006. "Value choices and American public opinion." *American Journal of Political Science* 50 (3): 706–723.
- Jensen, Jacob, Suresh Naidu, Ethan Kaplan, Laurence Wilse-Samson, David Gergen, Michael Zuckerman, and Arthur Spirling. 2012. "Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech [with comments and discussion]." *Brookings Papers on Economic Activity*: 1–81.
- Jerit, Jennifer, and Jason Barabas. 2017. "Revisiting the Gender Gap in Political Knowledge." *Political Behavior* 39 (4): 817–838.
- Jhaver, Shagun, Pranil Vora, and Amy Bruckman. 2017. Designing for Civil Conversations: Lessons Learned from ChangeMyView. Technical report Georgia Institute of Technology.
- Jost, John T. 2006. "The End of the End of Ideology." *American Psychologist* 61 (7): 651–670.
- Karpowitz, Christopher F, Tali Mendelberg, and Lee Shaker. 2012. "Gender inequality in deliberative participation." *American Political Science Review* 106 (03): 533–547.
- Kertzer, Joshua D., Kathleen E. Powers, Brian C. Rathbun, and Ravi Iyer. 2014. "Moral Support: How Moral Values Shape Foreign Policy Attitudes." *Journal of Politics* 76 (3): 825–840.
- Kidwell, Blair, Adam Farmer, and David M Hardesty. 2013. "Getting liberals and conservatives to go green: Political ideology and congruent appeals." *Journal of Consumer Research* 40 (2): 350–367.
- Klar, Samara. 2014. "Partisanship in a social setting." *American Journal of Political Science* 58 (3): 687–704.
- Koleva, Spassena P., Jesse Graham, Ravi Iyer, Peter H. Ditto, and Jonathan Haidt. 2012. "Tracing the Threads: How Five Moral Concerns (Especially Purity) Help Explain Culture War Attitudes." *Journal of Research in Personality* 46 (2): 184–194.

## CHAPTER 6. BIBLIOGRAPHY

---

- Kraft, Patrick W. 2018. "Measuring Morality in Political Attitude Expression." *The Journal of Politics* (forthcoming): <https://doi.org/10.1086/696862>.
- Krosnick, Jon A. 1999. "Survey research." *Annual review of psychology* 50 (1): 537–567.
- Kunda, Ziva, and Lisa Sinclair. 1999. "Motivated reasoning with stereotypes: Activation, application, and inhibition." *Psychological Inquiry* 10 (1): 12–22.
- Laver, Michael, and John Garry. 2000. "Estimating policy positions from political texts." *American Journal of Political Science*: 619–634.
- Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting policy positions from political texts using words as data." *American Political Science Review* 97 (02): 311–331.
- Lavine, Howard G, Christopher D Johnston, and Marco R Steenbergen. 2012. *The ambivalent partisan: How critical loyalty promotes democracy*. Oxford University Press.
- Lazer, David, Brian Rubineau, Carol Chetkovich, Nancy Katz, and Michael Neblo. 2010. "The coevolution of networks and political attitudes." *Political Communication* 27 (3): 248–274.
- Lee, Moontae, and David Mimno. 2014. "Low-dimensional embeddings for interpretable anchor-based topic inference." In *Proceedings of Empirical Methods in Natural Language Processing*. Citeseer.
- Lipsitz, Keena. 2018. "Playing with Emotions: The Effect of Moral Appeals in Elite Rhetoric." *Political Behavior* 40 (1): 57–78.
- Lizotte, Mary-Kate, and Andrew H Sidman. 2009. "Explaining the gender gap in political knowledge." *Politics & Gender* 5 (02): 127–151.
- Lodge, Milton, and Charles S Taber. 2013. *The Rationalizing Voter*. Cambridge University Press.
- Low, Michelle, and Ma Glenda Lopez Wui. 2015. "Moral Foundations and Attitudes Towards the Poor." *Current Psychology* 35: 650–656.
- Lupia, Arthur. 2006. "How elitism undermines the study of voter competence." *Critical Review* 18 (1-3): 217–232.
- Lupia, Arthur. 2015. *Uninformed: Why people seem to know so little about politics and what we can do about it*. Oxford University Press.
- Luskin, Robert C. 1987. "Measuring political sophistication." *American Journal of Political Science* 31 (4): 856–899.
- Manning, Christopher D., Prabhakar Raghavan, Hinrich Schütze et al. 2008. *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Mason, Lilliana. 2015. "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Behavioral and Issue Polarization." *American Journal of Political Science* 59 (1): 128–145.

- 
- McClurg, Scott D. 2006. "The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation." *American Journal of Political Science* 50 (3): 737–754.
- McClurg, S.D. 2003. "Social networks and political participation: The role of social interaction in explaining political participation." *Political Research Quarterly* 56 (4): 449–464.
- McDonald, John F., and Robert A. Moffitt. 1980. "The Uses of Tobit Analysis." *The Review of Economics and Statistics* 62 (2): 318–321.
- McGlone, Matthew S, Joshua Aronson, and Diane Kobrynowicz. 2006. "Stereotype threat and the gender gap in political knowledge." *Psychology of Women Quarterly* 30 (4): 392–398.
- Mendelberg, Tali, Christopher F Karpowitz, and J Baxter Oliphant. 2014. "Gender inequality in deliberation: Unpacking the black box of interaction." *Perspectives on Politics* 12 (01): 18–44.
- Mondak, Jeffery J, and Mary R Anderson. 2004. "The knowledge gap: A reexamination of gender-based differences in political knowledge." *Journal of Politics* 66 (2): 492–512.
- Mondak, Jeffrey J. 2001. "Developing valid knowledge scales." *American Journal of Political Science* 45 (1): 224–238.
- Mutz, Diana C. 2002. "Cross-cutting social networks: Testing democratic theory in practice." *American Political Science Review* 96 (01): 111–126.
- Nelson, Thomas E, and Jennifer Garst. 2005. "Values-based Political Messages and Persuasion: Relationships among Speaker, Recipient, and Evoked Values." *Political Psychology* 26 (4): 489–516.
- Nelson, Thomas E., Zoe M. Oxley, and Rosalee A. Clawson. 1997. "Toward a psychology of framing effects." *Political behavior* 19 (3): 221–246.
- Petty, Richard E, and John T Cacioppo. 1986a. *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer.
- Petty, Richard E, and John T Cacioppo. 1986b. "The elaboration likelihood model of persuasion." *Advances in experimental social psychology* 19: 123–205.
- Pietryka, Matthew T, and Randall C MacIntosh. 2013. "An Analysis of ANES Items and Their Use in the Construction of Political Knowledge Scales." *Political Analysis* 21 (4): 407–429.
- Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin, and Dragomir R Radev. 2010. "How to analyze political attention with minimal assumptions and costs." *American Journal of Political Science* 54 (1): 209–228.
- Redlawsk, David P. 2002. "Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making." *The Journal of Politics* 64 (04): 1021–1044.
- Roberts, Margaret E, Brandon M Stewart, and Dustin Tingley. 2014. "stm: R package for structural topic models." *Journal of Statistical Software* 1: 1–49.

## CHAPTER 6. BIBLIOGRAPHY

---

- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58 (4): 1064–1082.
- Rosenberg, Shawn W. 1988. "The structure of political thinking." *American Journal of Political Science*: 539–566.
- Rosenberg, Shawn W, Dana Ward, and Stephen Chilton. 1988. *Political reasoning and cognition: A Piagetian view*. Duke University Press.
- Ryan, John Barry. 2011. "Accuracy and bias in perceptions of political knowledge." *Political Behavior* 33 (2): 335–356.
- Ryan, Timothy J. 2014. "Reconsidering Moral Issues in Politics." *Journal of Politics* 76 (2): 380–397.
- Ryan, Timothy J. 2017. "No compromise: Political consequences of moralized attitudes." *American Journal of Political Science* 61 (2): 409–423.
- Skitka, Linda J. 2010. "The Psychology of Moral Conviction." *Social and Personality Psychology Compass* 4 (4): 267–281.
- Skitka, Linda J., Christopher W. Bauman, and Edward G. Sargis. 2005. "Moral Conviction: Another Contributor to Attitude Strength or Something More?" *Journal of Personality and Social Psychology* 88 (6): 895–917.
- Skitka, Linda J, and G Scott Morgan. 2014. "The social and political implications of moral conviction." *Political Psychology* 35 (S1): 95–110.
- Slapin, Jonathan B, and Sven-Oliver Proksch. 2008. "A scaling model for estimating time-series party positions from texts." *American Journal of Political Science* 52 (3): 705–722.
- Spirling, Arthur. 2016. "Democratization and Linguistic Complexity The Effect of Franchise Extension on Parliamentary Discourse, 1832–1915." *The Journal of Politics* 78 (1): 120–136.
- Stolle, Dietlind, and Elisabeth Gidengil. 2010. "What do women really know? A gendered analysis of varieties of political knowledge." *Perspectives on Politics* 8 (01): 93–109.
- Sturgis, Patrick, Nick Allum, and Patten Smith. 2008. "An experiment on the measurement of political knowledge in surveys." *Public Opinion Quarterly* 72 (1): 90–102.
- Sudman, Seymour, Norman M Bradburn, and Norbert Schwarz. 1996. *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.
- Tan, Chenhao, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. "Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions." In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee pp. 613–624.

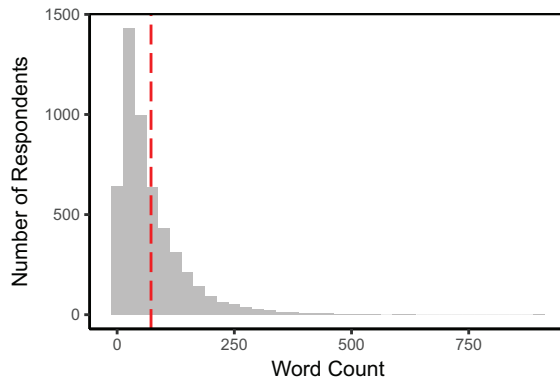
- 
- Tetlock, Philip E. 1983. "Cognitive style and political ideology." *Journal of Personality and Social Psychology* 45 (1): 118.
- Tetlock, Philip E. 1993. "Cognitive structural analysis of political rhetoric: Methodological and theoretical issues." In *Explorations in political psychology*. Duke University Press Durham.
- Verba, Sidney, Nancy Burns, and Kay Lehman Schlozman. 1997. "Knowing and caring about politics: Gender and political engagement." *The Journal of Politics* 59 (04): 1051–1072.
- Wei, Zhongyu, Yang Liu, and Yi Li. 2016. "Is this post persuasive? Ranking argumentative comments in online forum." In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Vol. 2 pp. 195–200.
- Wolak, Jennifer, and Michael McDevitt. 2011. "The roots of the gender gap in political knowledge in adolescence." *Political Behavior* 33 (3): 505–533.
- Zaller, John. 1990. "Political awareness, elite opinion leadership, and the mass survey response." *Social Cognition* 8 (1): 125.
- Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge University Press.

Appendix **A**

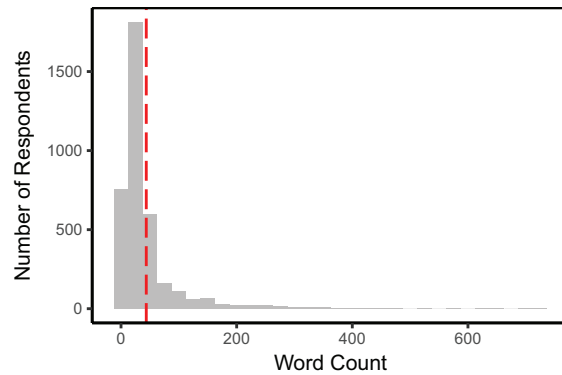
**Supplementary Material for Chapter 2**

## A.1 Detailed Information on Open-Ended Responses and Discursive Sophistication Components

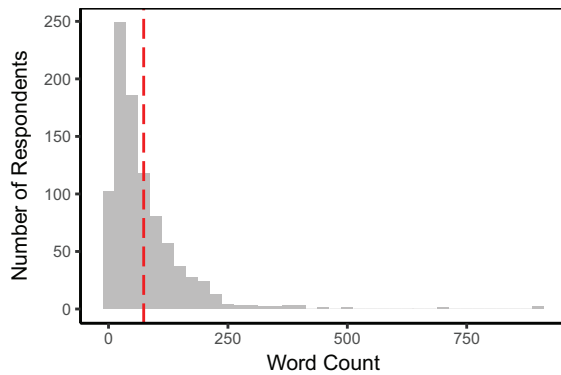
### A.1.1 Distribution of Word Counts in Open-Ended Responses



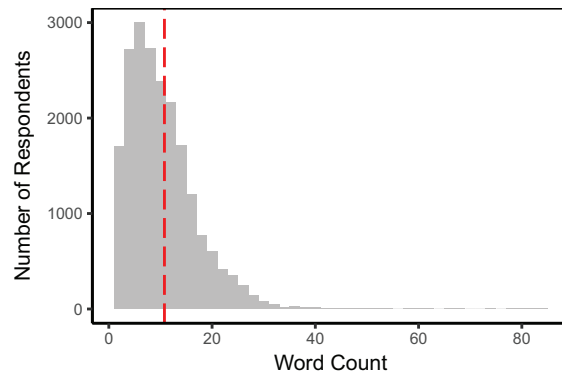
(a) 2012 ANES



(b) 2016 ANES



(c) 2015 YouGov



(d) Swiss Survey

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



A.1.2 Overview of Topic Proportions

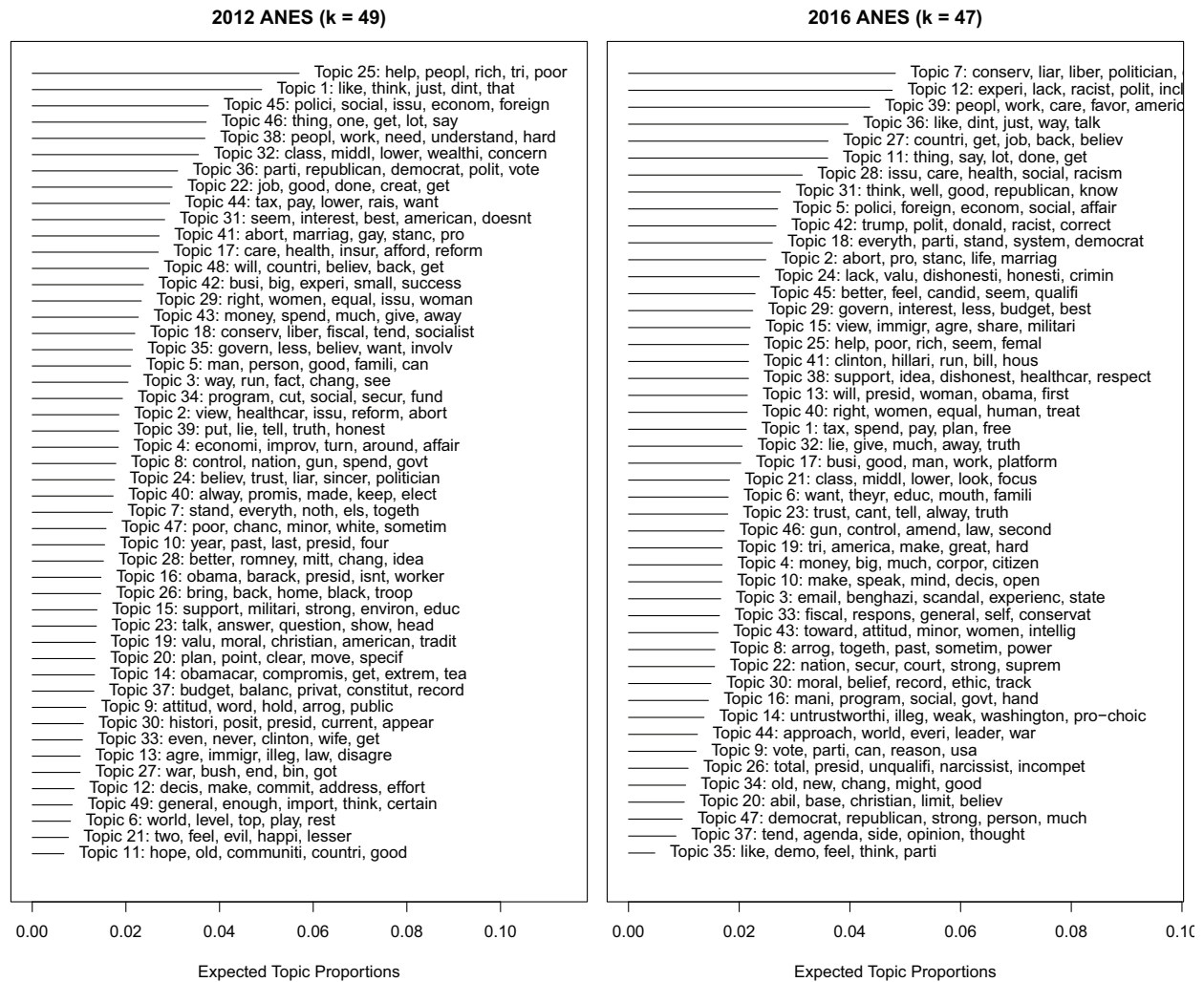


Figure A.2: Estimated topic proportions in the 2012 and 2016 ANES based on the structural topic model. See Appendix A.2 for details on the model specification.

## A.1. INFORMATION ON DISCURSIVE SOPHISTICATION COMPONENTS



Figure A.3: Estimated topic proportions in the 2015 YouGov survey based on the structural topic model. See Appendix A.2 for details on the model specification.

# APPENDIX A. SUPPLEMENTARY MATERIAL FOR CHAPTER 2

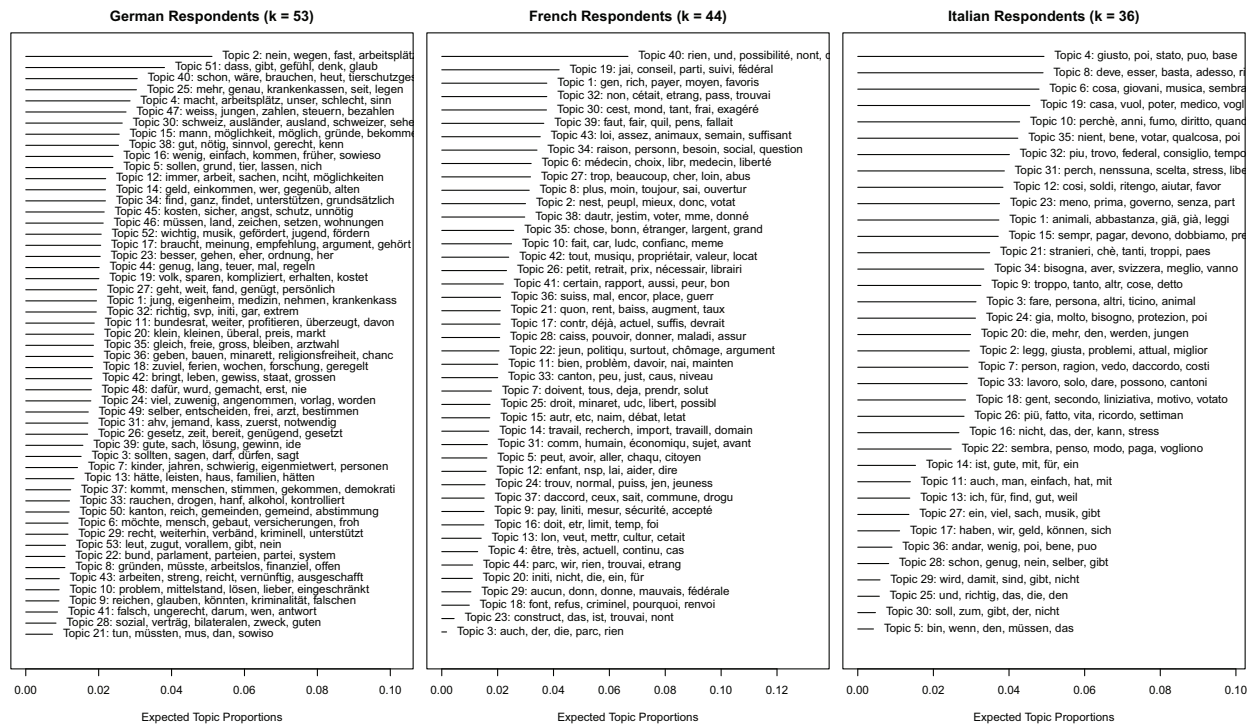
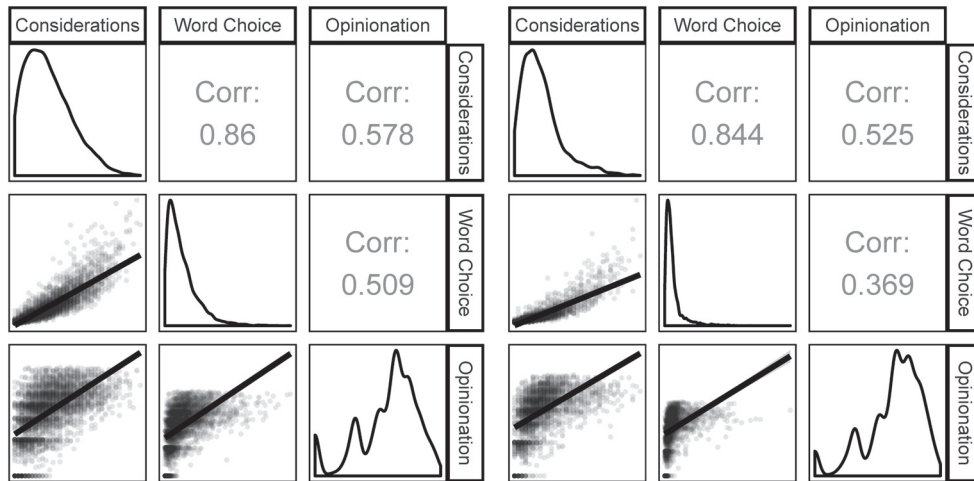


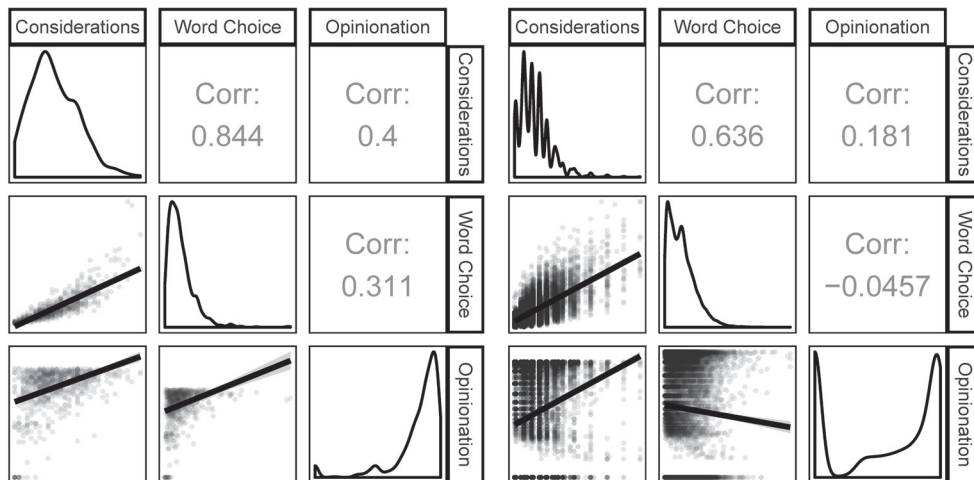
Figure A.4: Estimated topic proportions in the Swiss survey based on the structural topic model. See Appendix A.2 for details on the model specification.

A.1.3 Discursive Sophistication Components



(a) 2012 ANES

(b) 2016 ANES



(c) 2015 YouGov

(d) Swiss Survey

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

## A.2 Pre-Processing and Topic Model Specification

### A.2.1 PreText Analysis

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

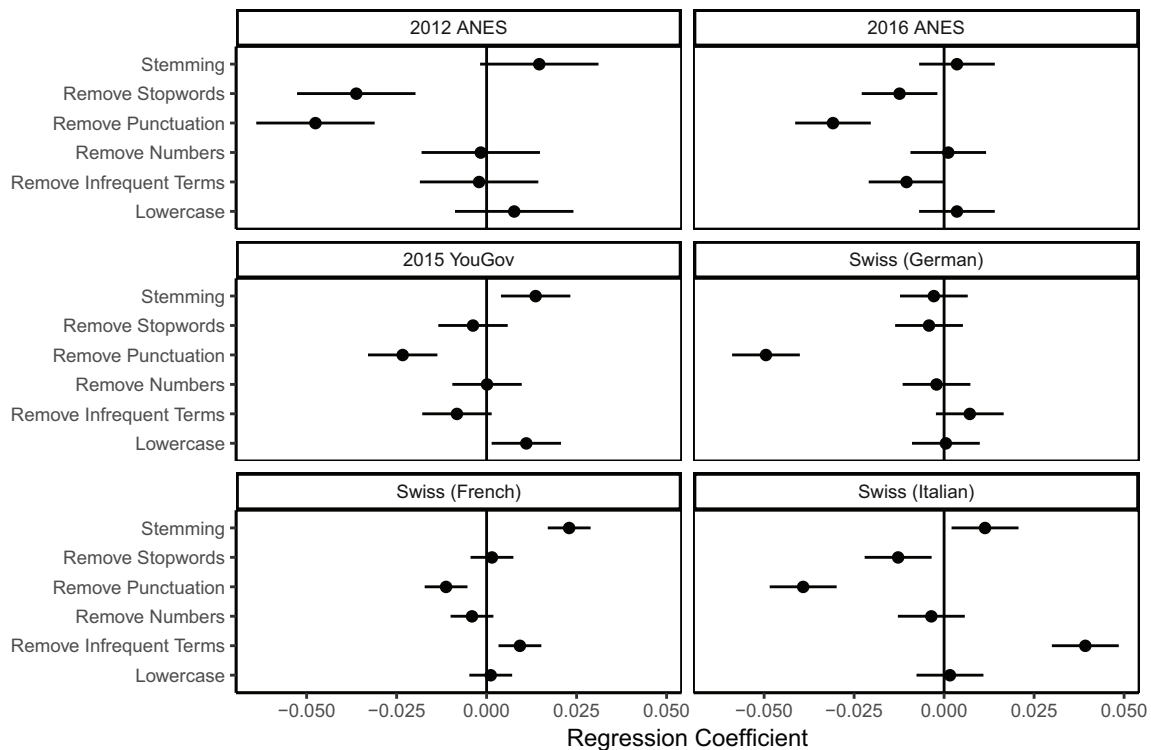


Figure A.6: PreText analysis of pre-processing decisions of open-ended responses across all datasets. Regression coefficients display the effects of each of the six pre-processing choices on the resulting preText score.

### A.2.2 Robustness Checks for Varying Model Specifications

Following the procedure outlined in Denny and Spirling (2018), Figure A.6 displays the results of a linear model regressing preText scores resulting from all possible pre-processing regimes on each individual step for a random subset of 500 open-ended responses in each of the studies included in the analyses. Significant coefficients indicate that the topic model results may be sensitive to the respective pre-processing step. As such, removing stopwords and punctuation, as well as removing infrequent terms (at least in the 2016 ANES) might be problematic. Denny and Spirling (2018), however, emphasize that the most important consideration in choosing pre-processing steps are theoretical. Given that the purpose of the topic model is to extract considerations related to political preferences, there are strong theoretical reasons to remove stopwords and punctuation from open-ended responses as they do not contain any relevant content. Furthermore, I apply lowercasing and stemming of terms to reduce resulting document term matrix to a computationally more manageable size and since these pre-processing steps should not be influential according to the preText analysis.

It is less obvious from a theoretical perspective whether to remove infrequent terms from open-ended responses, although it is preferred in order to make the estimation of the discursive sophistication components computationally efficient. Since the preText analysis for the 2016 ANES suggests that this pre-processing step might be influential, I compare discursive sophistication for both alternative regimes below (c.f., Denny and Spirling, 2018). Before turning to this sensitivity check, however, I consider another crucial modeling choice when working with topic models: determining the total number of topics  $k$  to be estimated. For all analyses reported below, the number of topics was selected using the algorithm proposed by Lee and Mimno (2014) and implemented in the `stm` package in **R** (Roberts, Stewart, and Tingley, 2014).<sup>1</sup>

Figure A.7 examines whether the proposed measure of discursive sophistication is sensitive to the removal of infrequent terms as well as the chosen number of topics  $k$ . The y-axis depicts the preferred pre-processing regime including all steps discussed above while the x-axis plots results for alternative specifications. The panels on the left compare the preferred specification to discursive sophistication based on a reduced number of topics ( $k = 20$ ). The middle panels additionally include infrequent terms instead of removing them.<sup>2</sup> The panels on the right do not perform stemming as part of the pre-processing step. Across all panels, discursive sophistication scores are highly correlated and therefore insensitive to pre-processing choices and varying numbers of topics.

In summary, open-ended responses in the analyses reported in the main text are pre-processed by stemming and lowercasing, as well as the removing numbers, punctuation, stopwords, and infrequent terms (i.e., terms that appear in fewer than 10 responses).<sup>3</sup> While the results discussed in the manuscript are based on this preferred specification, the substantive results are robust for alternative pre-processing regimes or varying numbers of topics.

<sup>1</sup>I used measures for age, gender, education, party identification, as well as an interaction between education and party identification as covariates for topic prevalence. This variable selection—with the exception of including gender—is equivalent to the procedure model specification described in Roberts et al. (2014).

<sup>2</sup>Calculating discursive sophistication with large numbers of topics while including infrequent terms is computationally prohibitive.

<sup>3</sup>Prior to applying these pre-processing steps, open-ended responses in the 2012 & 2016 ANES as well as the 2015 YouGov survey are cleaned by correcting spelling errors using an implementation of the Aspell spell-checking algorithm ([www.aspell.net](http://www.aspell.net)).



APPENDIX A. SUPPLEMENTARY MATERIAL FOR CHAPTER 2

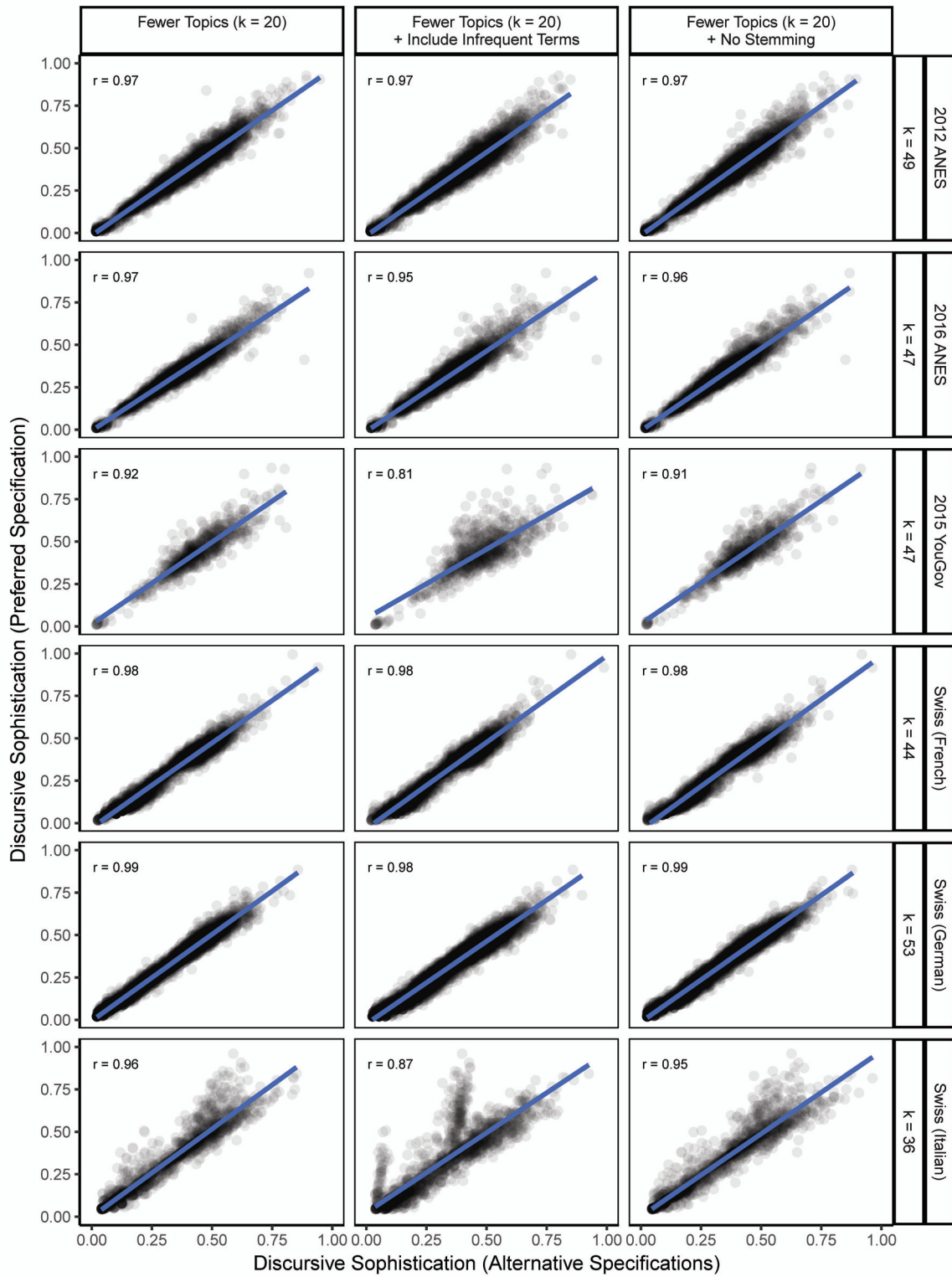


Figure A.7: Robustness of discursive sophistication measure for different pre-processing choices and topic model specifications.

### A.3 Additional Information on Remaining Variables

#### A.3.1 Item Selection and Recoding

##### Conventional measures of political knowledge:

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

##### Dependent variables:

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



### Control variables:

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

A.3.2 Variable Distributions – 2012 ANES

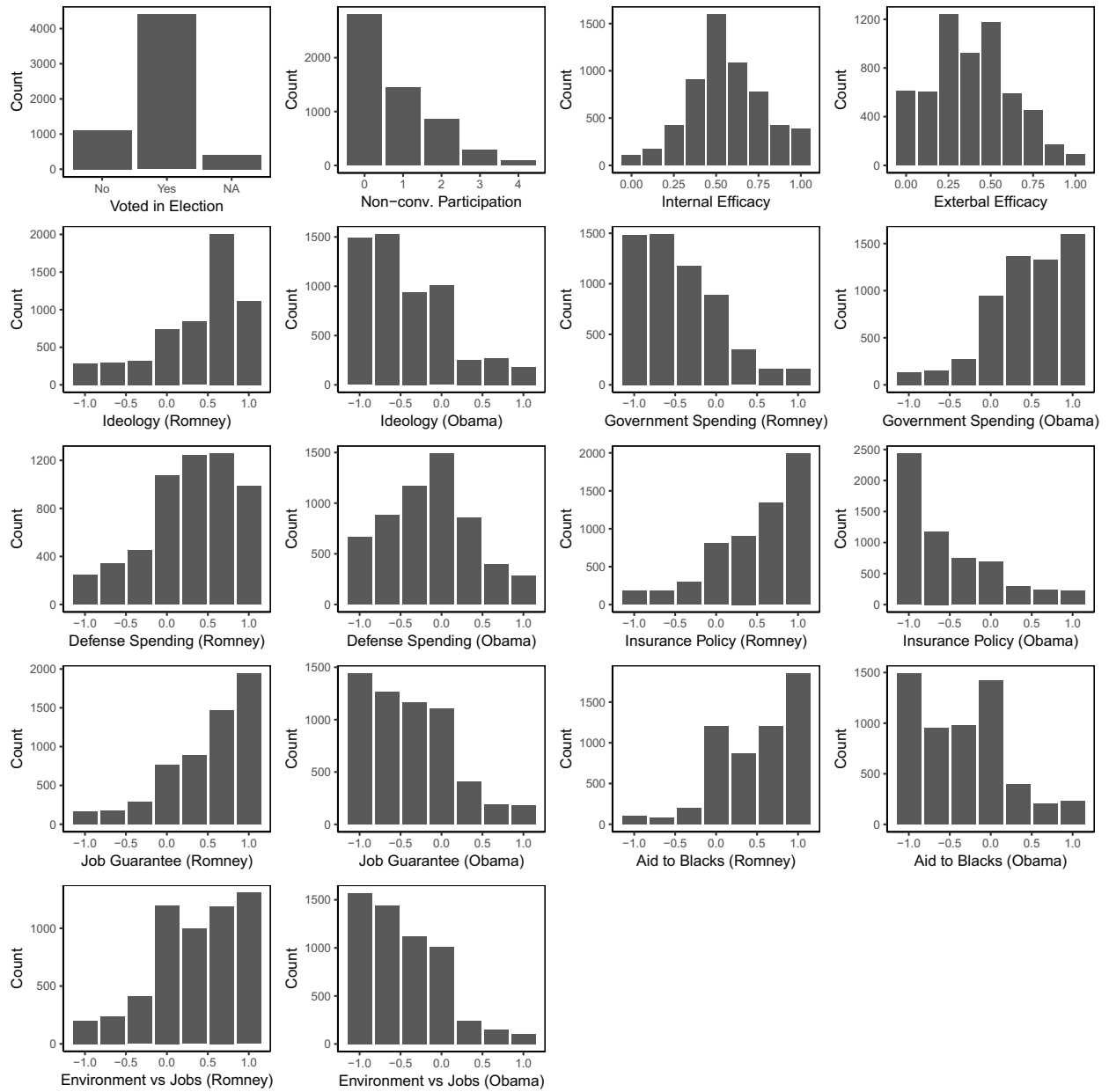


Figure A.8: Histograms of dependent variables included in 2012 ANES.

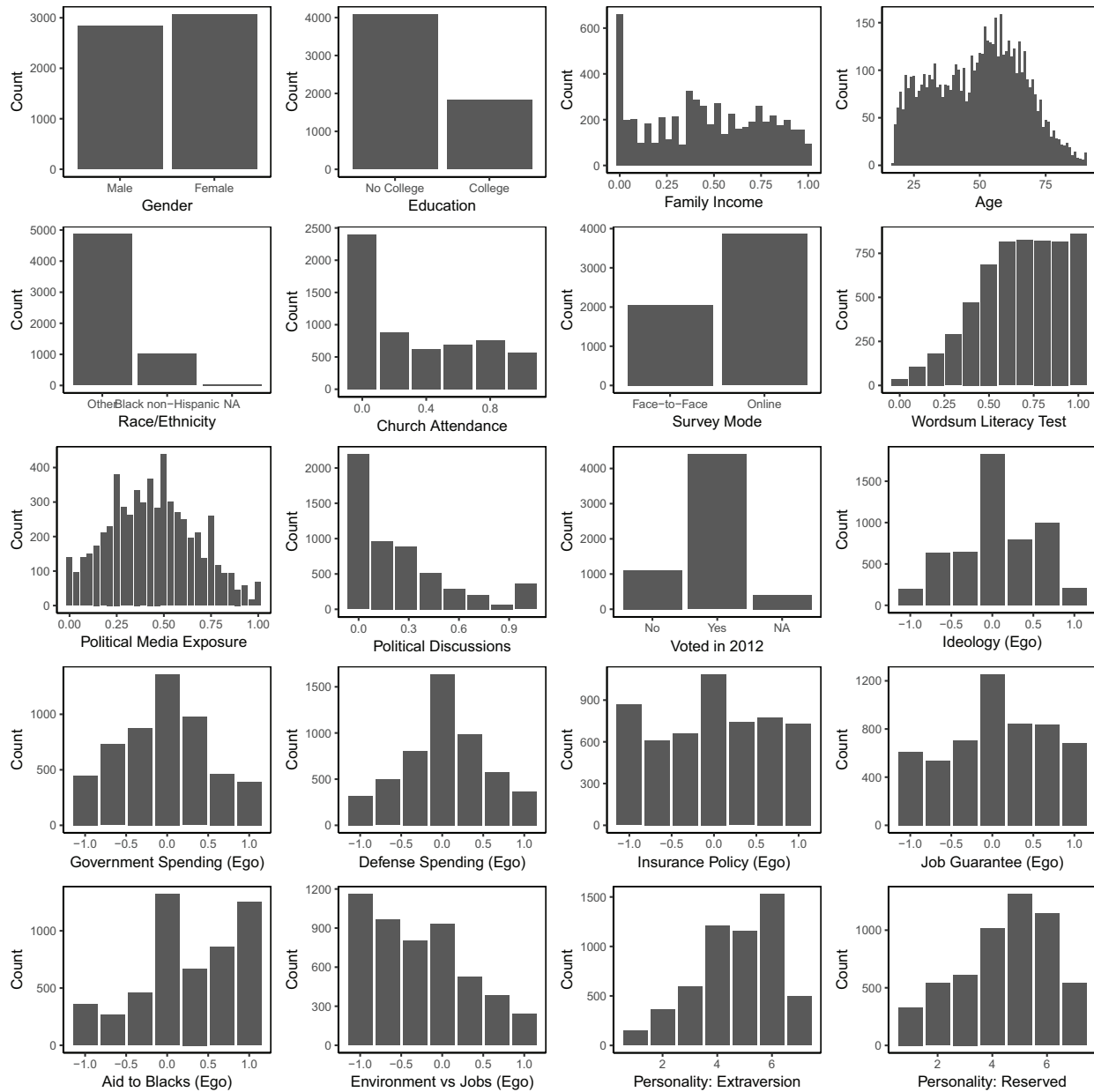


Figure A.9: Histograms of independent variables included in 2012 ANES.

A.3.3 Variable Distributions – 2016 ANES

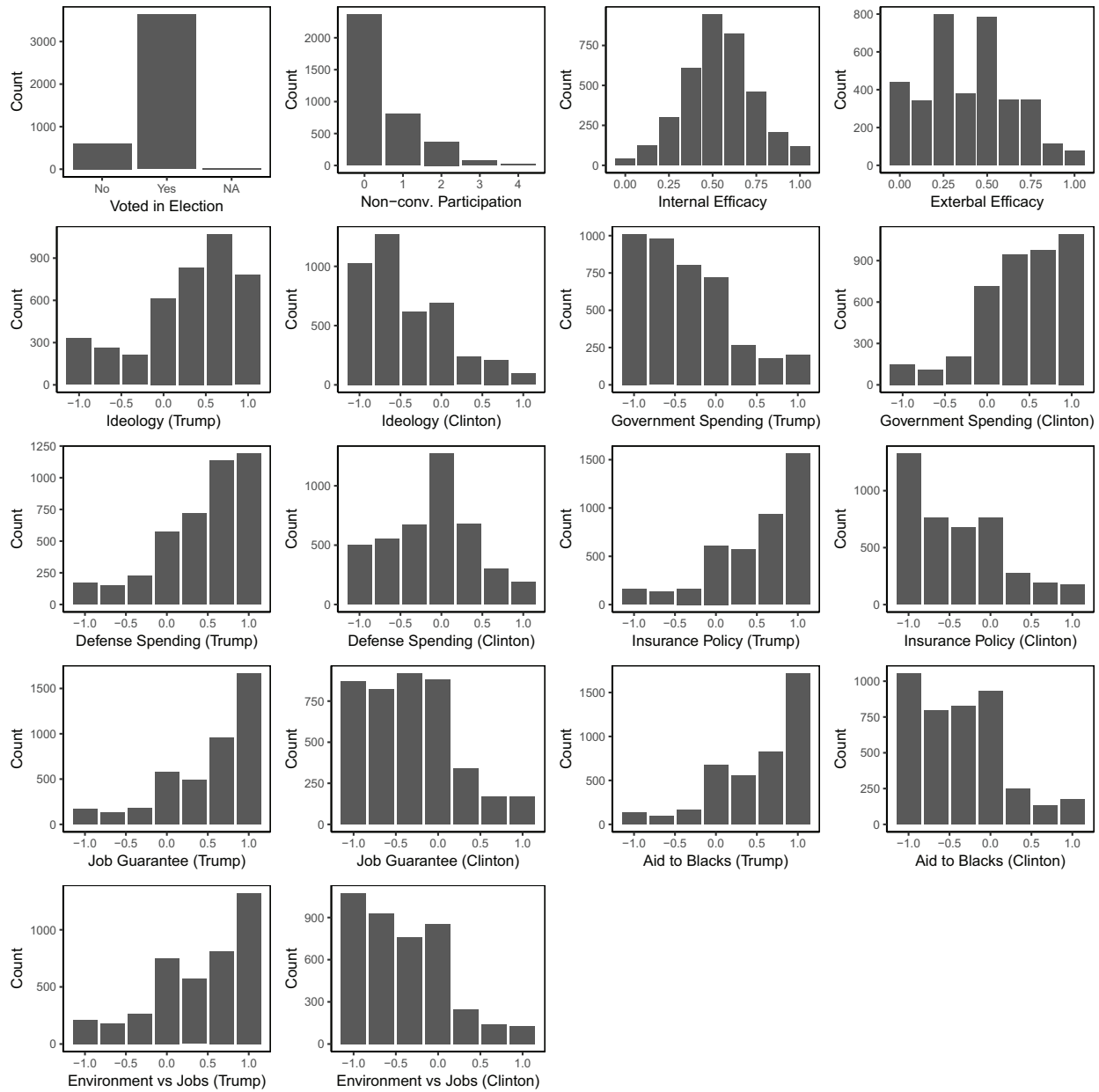


Figure A.10: Histograms of dependent variables included in 2016 ANES.

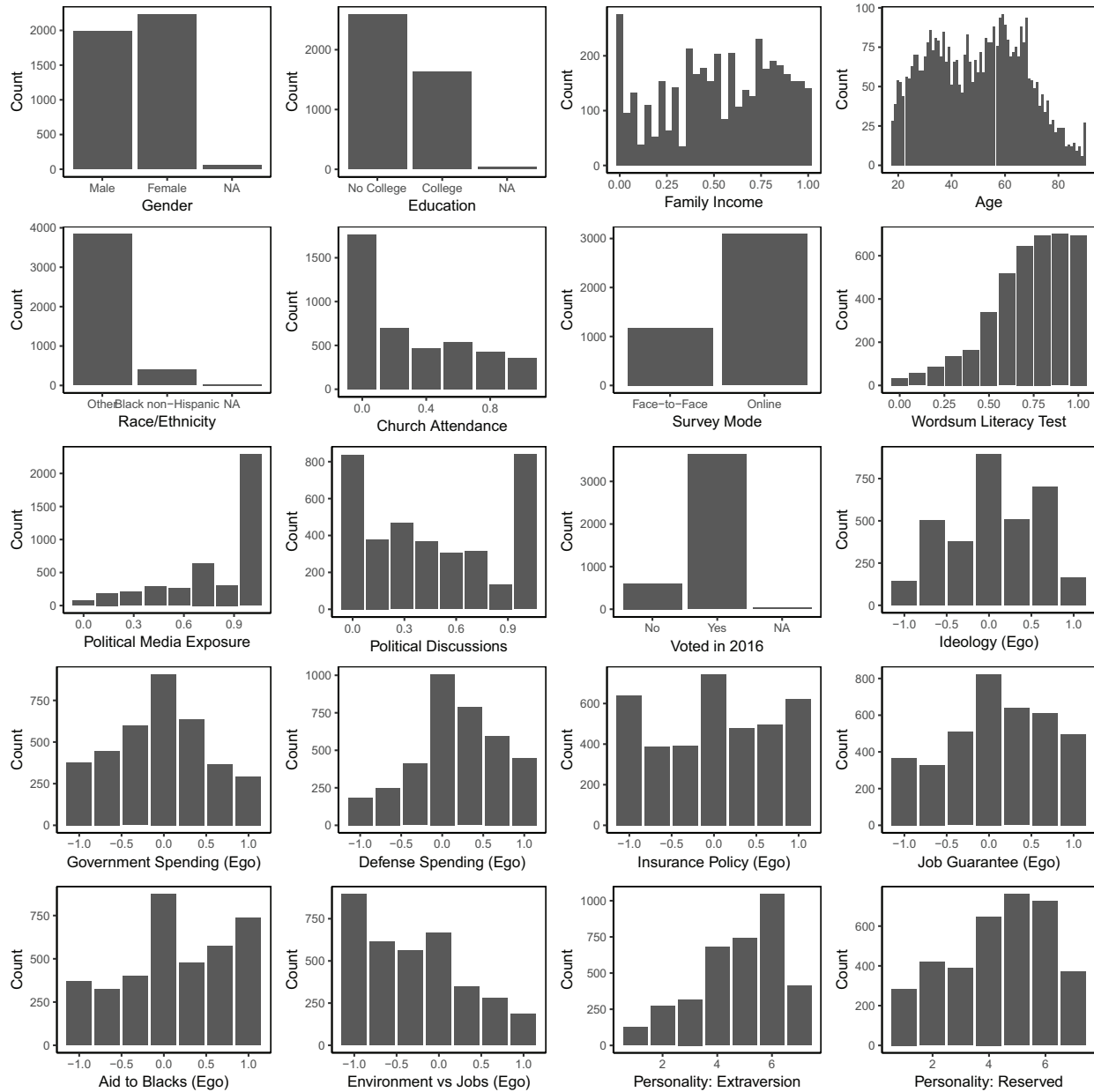


Figure A.11: Histograms of independent variables included in 2016 ANES.

A.3.4 Variable Distributions – 2015 YouGov

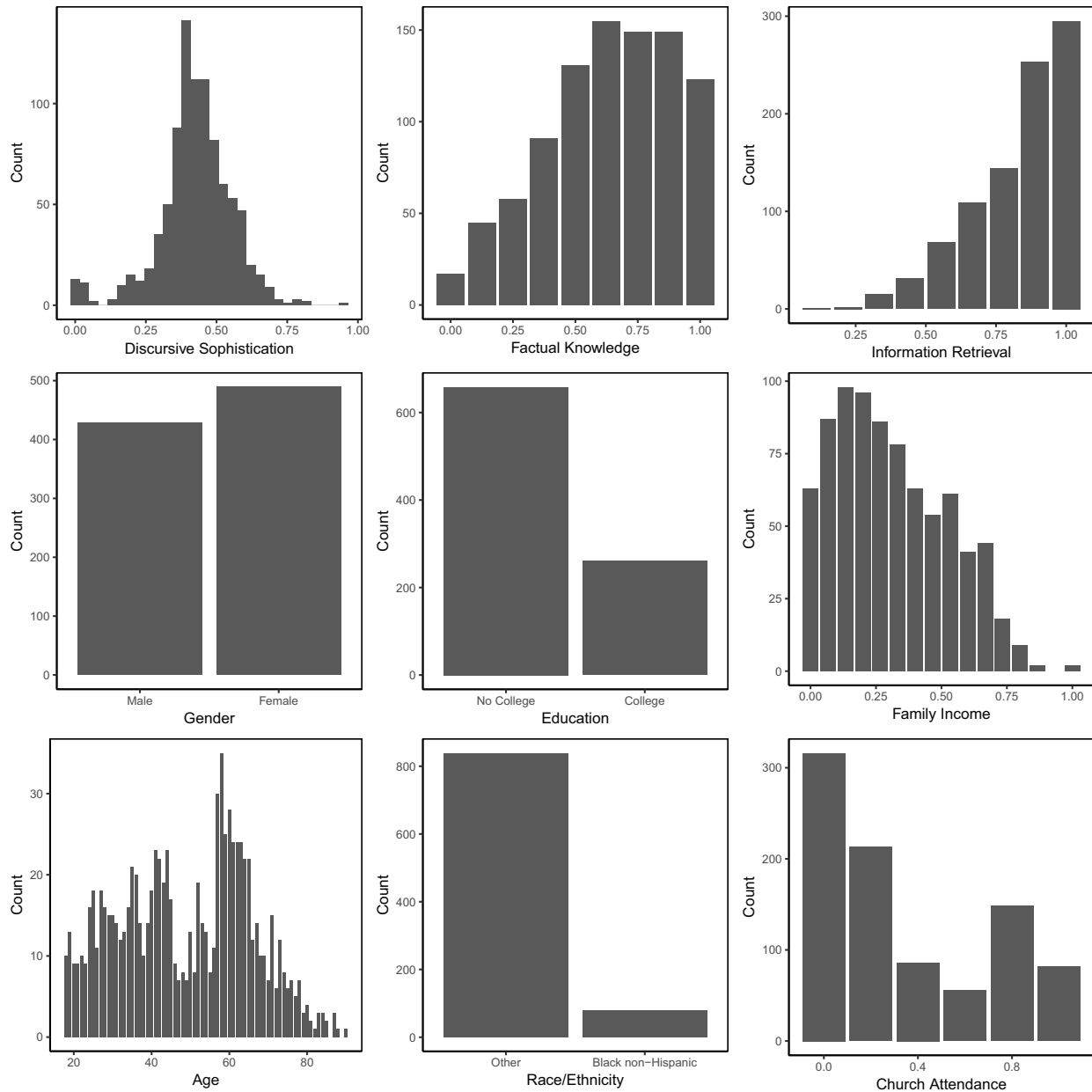


Figure A.12: Histogram of variables included in 2015 YouGov survey.

## A.4 Robustness Checks

### A.4.1 Controlling for Personality Characteristics

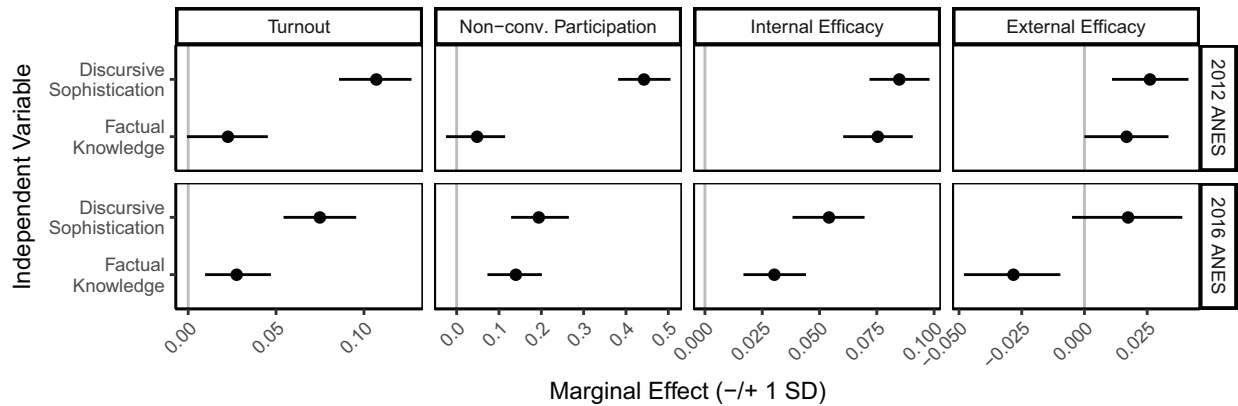


Figure A.13: Effects of sophistication on internal efficacy, external efficacy, non-conventional participation, and turnout in the 2012 and 2016 ANES. For each dependent variable, the figure displays the change in expected values after increasing each sophistication measure from -1 to +1 standard deviation from its mean (including 95% confidence intervals). Model estimates are based on logistic regression (turnout) or OLS (internal efficacy, external efficacy, non-conventional participation). Both sophistication measure are included simultaneously while controlling for gender, education, income, age, race, church attendance, survey mode, Wordsum vocabulary scores, as well as personality characteristics (extraversion and being reserved). Full model results are presented in the appendix, Tables A.8 and A.9.

### A.4.2 Controlling for Individual Response Length

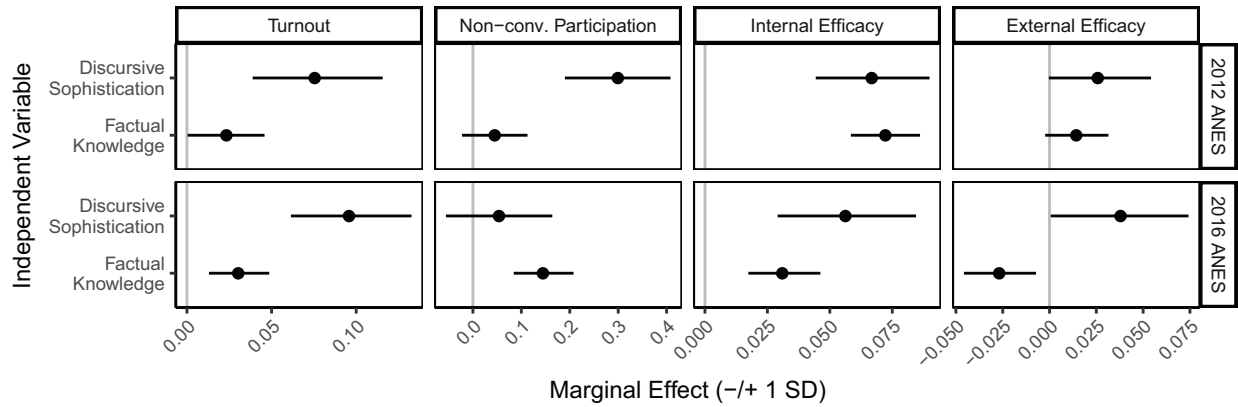


Figure A.14: Effects of sophistication on internal efficacy, external efficacy, non-conventional participation, and turnout in the 2012 and 2016 ANES. For each dependent variable, the figure displays the change in expected values after increasing each sophistication measure from -1 to +1 standard deviation from its mean (including 95% confidence intervals). Model estimates are based on logistic regression (turnout) or OLS (internal efficacy, external efficacy, non-conventional participation). Both sophistication measure are included simultaneously while controlling for gender, education, income, age, race, church attendance, survey mode, Wordsum vocabulary scores, as well as the logged word count across open-ended responses. Full model results are presented in the appendix, Tables A.10 and A.11.



## A.5 Tables of Model Estimates

### A.5.1 Main Analyses

Table A.1: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2012 ANES. Standard errors in parentheses. Estimates are used for Figure 2.2 in the main text.

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

## A.5. TABLES OF MODEL ESTIMATES

Table A.2: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2016 ANES. Standard errors in parentheses. Estimates are used for Figure 2.2 in the main text.

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

Table A.3: Effects of sophistication on information retrieval in the 2015 YouGov study. Standard errors in parentheses. Estimates are used for Figure 2.3 in the main text.

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

## A.5. TABLES OF MODEL ESTIMATES

Table A.4: Error variance reduction in candidate placements on multiple issues in the 2012 ANES. Estimates are used for Figure 2.4 in the main text.

Policy Position	Candidate	Independent Var.	$E[\gamma]$	$sd(\gamma)$	95% Cred. Int.	$\hat{R}$
Ideology	Romney	Discursive Soph.	-0.71	0.07	(-0.856; -0.562)	1.00
Ideology	Romney	Factual Knowl.	-0.65	0.06	(-0.755; -0.538)	1.00
Ideology	Obama	Discursive Soph.	-0.41	0.07	(-0.544; -0.264)	1.00
Ideology	Obama	Factual Knowl.	-0.58	0.06	(-0.69; -0.47)	1.00
Government Spending	Romney	Discursive Soph.	-0.34	0.07	(-0.482; -0.205)	1.00
Government Spending	Romney	Factual Knowl.	-0.41	0.06	(-0.52; -0.303)	1.00
Government Spending	Obama	Discursive Soph.	-0.54	0.07	(-0.68; -0.398)	1.00
Government Spending	Obama	Factual Knowl.	-0.47	0.06	(-0.582; -0.354)	1.00
Defense Spending	Romney	Discursive Soph.	-0.18	0.08	(-0.326; -0.031)	1.00
Defense Spending	Romney	Factual Knowl.	-0.33	0.06	(-0.443; -0.223)	1.00
Defense Spending	Obama	Discursive Soph.	-0.17	0.08	(-0.318; -0.028)	1.00
Defense Spending	Obama	Factual Knowl.	-0.19	0.05	(-0.287; -0.086)	1.00
Insurance Policy	Romney	Discursive Soph.	-0.26	0.07	(-0.398; -0.12)	1.00
Insurance Policy	Romney	Factual Knowl.	-0.22	0.05	(-0.326; -0.116)	1.00
Insurance Policy	Obama	Discursive Soph.	-0.33	0.07	(-0.476; -0.194)	1.00
Insurance Policy	Obama	Factual Knowl.	-0.41	0.06	(-0.515; -0.297)	1.00
Job Guarantee	Romney	Discursive Soph.	-0.59	0.08	(-0.734; -0.436)	1.00
Job Guarantee	Romney	Factual Knowl.	-0.33	0.06	(-0.439; -0.223)	1.00
Job Guarantee	Obama	Discursive Soph.	-0.35	0.07	(-0.498; -0.215)	1.00
Job Guarantee	Obama	Factual Knowl.	-0.20	0.05	(-0.304; -0.091)	1.00
Aid to Blacks	Romney	Discursive Soph.	-0.22	0.07	(-0.368; -0.069)	1.00
Aid to Blacks	Romney	Factual Knowl.	-0.27	0.05	(-0.369; -0.163)	1.00
Aid to Blacks	Obama	Discursive Soph.	-0.21	0.07	(-0.36; -0.07)	1.00
Aid to Blacks	Obama	Factual Knowl.	-0.20	0.06	(-0.312; -0.095)	1.00
Environment vs Jobs	Romney	Discursive Soph.	-0.29	0.08	(-0.438; -0.138)	1.00
Environment vs Jobs	Romney	Factual Knowl.	-0.34	0.06	(-0.456; -0.227)	1.00
Environment vs Jobs	Obama	Discursive Soph.	-0.34	0.08	(-0.484; -0.181)	1.00
Environment vs Jobs	Obama	Factual Knowl.	-0.35	0.06	(-0.465; -0.233)	1.00

## APPENDIX A. SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Table A.5: Error variance reduction in candidate placements on multiple issues in the 2016 ANES. Estimates are used for Figure 2.4 in the main text.

Policy Position	Candidate	Independent Var.	$E[\gamma]$	$sd(\gamma)$	95% Cred. Int.	$\hat{R}$
Ideology	Trump	Discursive Soph.	-0.77	0.11	(-0.984; -0.55)	1.00
Ideology	Trump	Factual Knowl.	-0.16	0.06	(-0.272; -0.057)	1.00
Ideology	Clinton	Discursive Soph.	-0.52	0.10	(-0.71; -0.313)	1.00
Ideology	Clinton	Factual Knowl.	-0.15	0.05	(-0.255; -0.048)	1.00
Government Spending	Trump	Discursive Soph.	-0.62	0.10	(-0.835; -0.424)	1.00
Government Spending	Trump	Factual Knowl.	-0.09	0.05	(-0.19; 0.013)	1.00
Government Spending	Clinton	Discursive Soph.	-1.00	0.11	(-1.213; -0.785)	1.00
Government Spending	Clinton	Factual Knowl.	-0.18	0.05	(-0.284; -0.079)	1.00
Defense Spending	Trump	Discursive Soph.	-0.37	0.10	(-0.563; -0.171)	1.00
Defense Spending	Trump	Factual Knowl.	-0.08	0.05	(-0.183; 0.021)	1.00
Defense Spending	Clinton	Discursive Soph.	-0.54	0.11	(-0.744; -0.319)	1.00
Defense Spending	Clinton	Factual Knowl.	-0.04	0.05	(-0.141; 0.061)	1.00
Insurance Policy	Trump	Discursive Soph.	-0.79	0.10	(-0.991; -0.579)	1.00
Insurance Policy	Trump	Factual Knowl.	-0.06	0.05	(-0.168; 0.04)	1.00
Insurance Policy	Clinton	Discursive Soph.	-0.74	0.11	(-0.959; -0.521)	1.00
Insurance Policy	Clinton	Factual Knowl.	-0.14	0.05	(-0.241; -0.033)	1.00
Job Guarantee	Trump	Discursive Soph.	-0.86	0.10	(-1.061; -0.671)	1.00
Job Guarantee	Trump	Factual Knowl.	-0.06	0.05	(-0.157; 0.043)	1.00
Job Guarantee	Clinton	Discursive Soph.	-0.76	0.10	(-0.965; -0.546)	1.00
Job Guarantee	Clinton	Factual Knowl.	-0.04	0.05	(-0.134; 0.059)	1.00
Aid to Blacks	Trump	Discursive Soph.	-0.66	0.10	(-0.862; -0.466)	1.00
Aid to Blacks	Trump	Factual Knowl.	-0.00	0.05	(-0.099; 0.097)	1.00
Aid to Blacks	Clinton	Discursive Soph.	-0.47	0.10	(-0.663; -0.277)	1.00
Aid to Blacks	Clinton	Factual Knowl.	-0.06	0.05	(-0.158; 0.045)	1.00
Environment vs Jobs	Trump	Discursive Soph.	-0.65	0.11	(-0.861; -0.443)	1.00
Environment vs Jobs	Trump	Factual Knowl.	-0.18	0.05	(-0.281; -0.072)	1.00
Environment vs Jobs	Clinton	Discursive Soph.	-0.76	0.10	(-0.956; -0.554)	1.00
Environment vs Jobs	Clinton	Factual Knowl.	-0.24	0.05	(-0.348; -0.142)	1.00

Table A.6: Effects of gender on discursive sophistication and factual knowledge in the 2012 ANES and 2016 ANES. Standard errors in parentheses. Estimates are used for Figure 2.7 in the main text.

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

Table A.7: Effects of gender on discursive sophistication and factual knowledge in the 2015 YouGov Study. Standard errors in parentheses. Estimates are used for Figure 2.7 in the main text.

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

## A.5.2 Robustness Checks

Table A.8: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for personality characteristics in the 2012 ANES. Standard errors in parentheses. Estimates are used for Figure A.13 in the appendix.

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



Table A.9: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for personality characteristics in the 2016 ANES. Standard errors in parentheses. Estimates are used for Figure A.13 in the appendix.

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

## A.5. TABLES OF MODEL ESTIMATES

Table A.10: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for individual response length in the 2012 ANES. Standard errors in parentheses. Estimates are used for Figure A.14 in the appendix.

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

Table A.11: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for individual response length in the 2016 ANES. Standard errors in parentheses. Estimates are used for Figure A.14 in the appendix.

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

Appendix **B**

**Supplementary Material for Chapter 3**

## B.1 Moral Foundations Dictionary

### Sources:

Graham, Haidt, and Nosek (2009), as well as <http://www.moralfoundations.org/>

### Note:

Terms with (\*) indicate that the word stem rather than the exact word was matched in the open-ended survey responses.

### Care:

safe\*, peace\*, compassion\*, empath\*, sympath\*, care, caring, protect\*, shield, shelter, amity, secur\*, benefit\*, defen\*, guard\*, preserve, harm\*, suffer\*, war, wars, warl\*, warring, fight\*, violen\*, hurt\*, kill, kills, killer\*, killed, killing, endanger\*, cruel\*, brutal\*, abuse\*, damag\*, ruin\*, ravage, detriment\*, crush\*, attack\*, annihilate\*, destroy, stomp, abandon\*, spurn, impair, exploit, exploits, exploited, exploiting, wound\*

### Fairness:

fair, fairly, fairness, fair\*, fairmind\*, fairplay, equal\*, justice, justness, justifi\*, reciproc\*, impartial\*, egalitar\*, rights, equity, evenness, equivalent, unbias\*, tolerant, equable, balance\*, homologous, unprejudice\*, reasonable, constant, honest\*, unfair\*, unequal\*, bias\*, unjust\*, injust\*, bigot\*, discriminat\*, disproportion\*, inequitable, prejud\*, dishonest, unscrupulous, dissociate, preference, favoritism, segregat\*, exclusion, exclud\*

### Loyalty:

together, nation\*, homeland\*, family, families, familial, group, loyal\*, patriot\*, communal, commune\*, communit\*, communis\*, comrad\*, cadre, collectiv\*, joint, unison, unite\*, fellow\*, guild, solidarity, devot\*, member, cliqu\*, cohort, ally, insider, foreign\*, enem\*, betray\*, treason\*, traitor\*, treacher\*, disloyal\*, individual\*, apostasy, apostate, deserted, deserter\*, deserting, deceiv\*, jilt\*, imposter, miscreant, spy, sequester, renegade, terroris\*, immigra\*

### Authority:

obey\*, obedien\*, duty, law, lawful\*, legal\*, duti\*, honor\*, respect, respectful\*, respected, respects, order\*, father\*, mother, motherl\*, mothering, mothers, tradition\*, hierarch\*, authorit\*, permit, permission, status\*, rank\*, leader\*, class, bourgeoisie, caste\*, position, complian\*, command, supremacy, control, submi\*, allegian\*, serve, abide, defere\*, defer, revere\*, venerat\*, comply, defian\*, rebel\*, dissent\*, subver\*, disrespect\*, disobe\*, sediti\*, agit\*at\*, insubordinat\*, illegal\*, lawless\*, insurgent, mutinous, defy\*, dissident, unfaithful, alienate, defector, heretic\*, nonconformist, oppose, protest, refuse, denounce, remonstrate, riot\*, obstruct

### Sanctity:

piety, pious, purity, pure\*, clean\*, steril\*, sacred\*, chast\*, holy, holiness, saint\*, wholesome\*, celiba\*, abstention, virgin, virgins, virginity, virginal, austerity, integrity, modesty, abstin\*en\*, abstemiousness, upright, limpid, unadulterated, maiden, virtuous, refined, intemperate, decen\*, immaculate, innocent, pristine, humble, disgust\*, deprav\*, disease\*, unclean\*, contagio\*

## B.1. MORAL FOUNDATIONS DICTIONARY

---

indecen\*, sin, sinful\*, sinner\*, sins, sinned, sinning, slut\*, whore, dirt\*, impiety, impious, profan\*, gross, repuls\*, sick\*, promiscu\*, lewd\*, adulter\*, debauch\*, defile\*, tramp, prostitut\*, unchaste, wanton, profligate, filth\*, trashy, obscen\*, lax, taint\*, stain\*, tarnish\*, debase\*, desecrat\*, wicked\*, blemish, exploitat\*, pervert, wretched\*

## B.2 Information on Data, Variables, and Recoding

### B.2.1 Open-ended Responses and MFT Scores in 2012 ANES

In this study, MFT scores are computed based on verbatim open-ended responses in which individuals describe what they *liked* and *disliked* about either presidential candidate as well as the Republican and Democratic parties. More specifically, respondents in the 2012 American National Election Study (ANES) were asked to list anything in particular that they like/dislike about the Democratic/Republican party as well as anything that might make them vote/not vote for either of the Presidential candidates and were probed by the interviewer asking “anything else?” until the respondent answered “no.” All responses to the eight open-ended like/dislike questions (evaluating both parties and both candidates) were combined for each individual and pre-processed by correcting spelling errors.<sup>1</sup>

Table B.1: Missing open-ended responses

	N	Percent
Spanish Interview	228	3.86
No/Short Responses	655	11.08

Respondents were not included in the analysis if they only provided extremely short responses (less or equal to 5 words across all 8 open-ended items), or if the interview language was Spanish. Table B.1 provides an overview of the number of omitted cases. About 4% of the interviews were held in Spanish and about 11% of the respondents only provided extremely short or no responses (7% did not answer any open-ended question). Furthermore, Figure B.1 displays histograms of the length of the respondents’ answers to all open-ended items. On average, the collection of all open-ended responses consists of about 75 words for each individual.

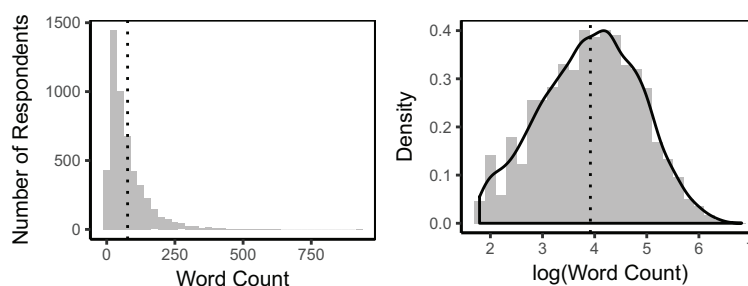


Figure B.1: Histograms displaying the distribution of individual response lengths in number of words for each respective item category. Dotted lines indicate the average response length.

Figure B.2 presents the proportion of respondents who mentioned words that were included in the five different moral foundations dictionaries. Since responses for each individual represent their likes and dislikes across all eight open-ended items, each proportion indicates the percentage

<sup>1</sup>Spell checks and corrections were implemented using GNU Aspell, see [www.aspell.net](http://www.aspell.net).

## B.2. INFORMATION ON DATA, VARIABLES, AND RECODING

of individuals who mentioned a signal word belonging to the respective moral foundation in any of his or her open-ended responses evaluating the parties or candidates.

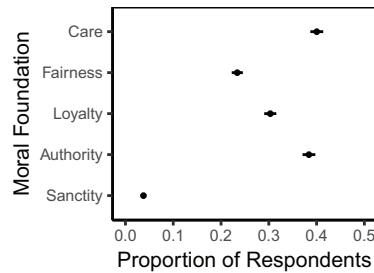


Figure B.2: Proportion of respondents mentioning each of the moral foundations in any of their open-ended responses, along with 95% confidence intervals. The first two foundations are often labeled individualizing foundations, which have been shown to be more prevalent among liberals, while the remaining ones are described as binding foundations, which are more prevalent among conservatives.

The moral foundation most frequently mentioned is *care*: About 40% of the respondents mentioned at least one word included in respective dictionary. The second most frequently mentioned moral foundation is *authority* with about 38%. The proportion of respondents emphasizing *loyalty* or *fairness* is slightly lower with about 30% and 23%, respectively. *Sanctity*, on the other hand, was almost never mentioned by any of the respondents, which suggests that the terms contained in the sanctity dictionary might be too uncommon in the context of politics and therefore not relevant for attitude expression. Due to the very rare mentioning of the sanctity dimension, the analyses in the main text concentrate on the remaining four moral foundations.<sup>2</sup> Subsequent analyses focusing on the sanctity dimension in open-ended survey responses might necessitate a revision of the moral foundations dictionary. Overall, Figure B.2 shows that a substantial proportion of individuals evokes moral considerations when describing their political attitudes even when they are not explicitly asked about morality.

The article proposes a weighting method to improve conventional dictionary approaches in order to capture the relative emphasis on each moral dimension (while correcting for ubiquitous terms and overall response length). Figure B.3 displays the weights used for each dictionary term that is mentioned at least once (terms that never appear are omitted). Terms that are very common, like “care”, or “foreign”, receive comparatively low weights. Such common words are more likely to be used in multiple contexts and cannot be uniquely ascribed to the moral domain, while dictionary terms that only appear in few responses are more likely to signal moral reasoning.

In parts of the analyses, MFT scores are combined to measure the overall reliance on moral considerations. This variable is computed as the sum of individual MFT scores across all dimensions (rescaled to unit variance after summation), which can be interpreted as a measure of general moralization in attitude expression.

<sup>2</sup>Unfortunately, this issue cannot not be addressed by relying on weighting scheme proposed in this study. The weights can correct for some distortions due to individual ubiquitous terms in the dictionaries, but it cannot compensate for the fact that the sanctity dictionary as a whole contains mostly words that are never mentioned by respondents.



## APPENDIX B. SUPPLEMENTARY MATERIAL FOR CHAPTER 3

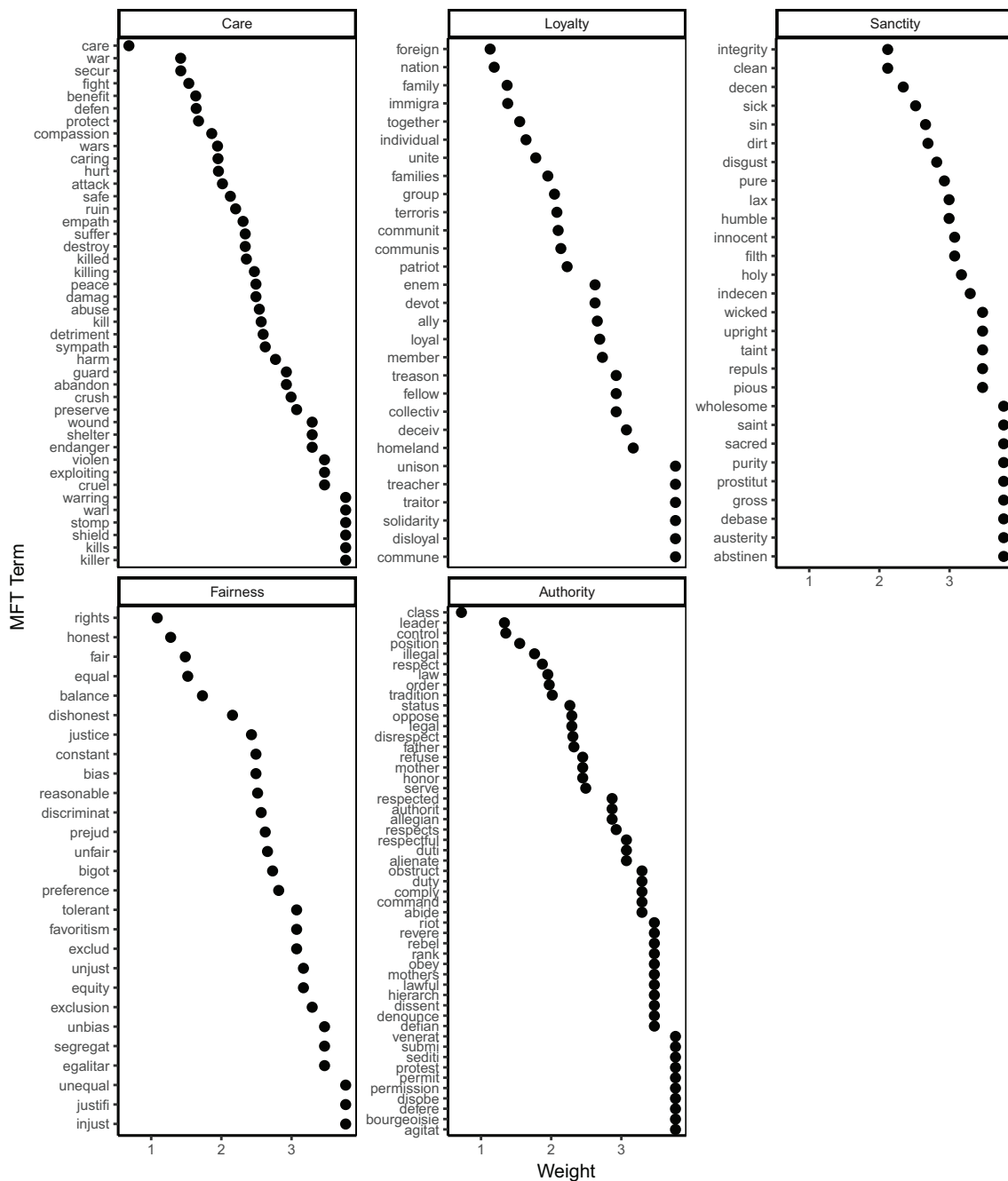


Figure B.3: Weights for individual MFT dictionary terms (terms that were not mentioned by any respondent are excluded).

### B.2.2 Moralization in Individual Media Environments

The last analysis in the main text explores the effect of individual media environments on general moralization in political attitude expression. I make use of the fact that the 2012 ANES included a large array of items indicating whether individuals regularly watched various news outlets. For all media sources available, I downloaded the content of the coverage on either presidential candidates during survey period in the last month of the campaign (October 2012) from Lexis-Nexis and computed their respective general MFT scores. Figure B.4 displays the results for each media outlet under consideration.

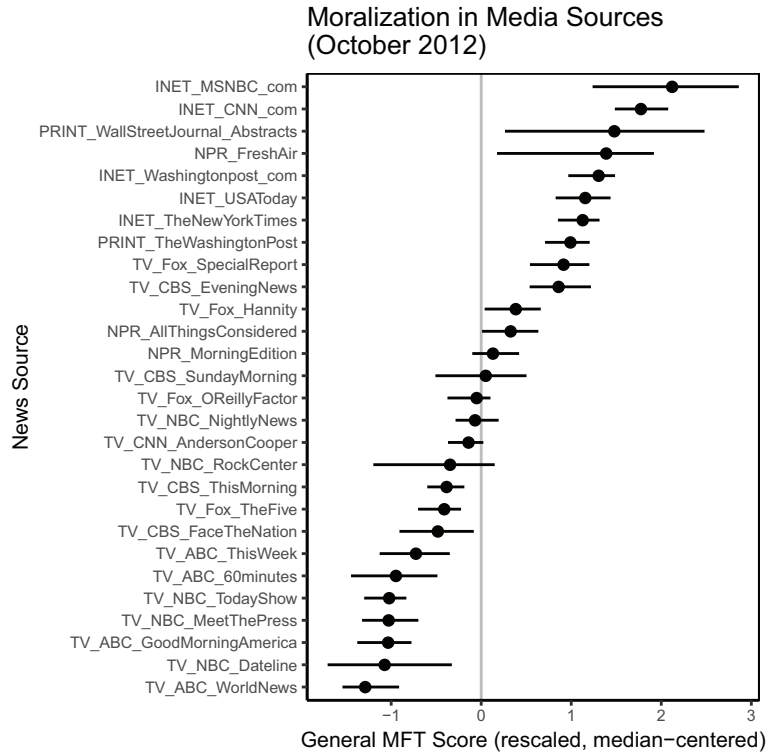


Figure B.4: General MFT scores for media sources during 2012 U.S. Presidential campaign. Articles and scripts were selected if they mentioned either presidential candidate during the survey field period in the last month of the campaign (October). Contents were retrieved in full text from Lexis-Nexis (except for the Wall Street Journal, which only provided abstracts). Each media source was analyzed using the same procedure described for open-ended responses (general moralization). Resulting general MFT scores were median-centered to facilitate the interpretation of relative moralization in media environments. The figure also displays 95% confidence intervals, which are based on parametric bootstraps of the document feature matrix of the entire corpus (1000 iterations).

Based on the coded content for each media source, I created a measure that represents the extent to which each individual’s media environment emphasized any moral foundation. For each respondent in the ANES, I select the media sources he or she reported to watch/read regularly and computed the mean of their respective general MFT scores. Figure B.5 displays the histogram

of the resulting variable. Negative values indicate that the average news source consumed by a respondent scores below the median in general moralization, while positive values indicate that the average news source consumed by a respondent emphasized moral foundations more strongly. Individuals who did not report to have watched or read any of the media outlets were omitted from this analysis. Using this approach, I can analyze whether individuals who rely on media sources that emphasize moral foundations were also more likely to mention the moral considerations in their open-ended responses.

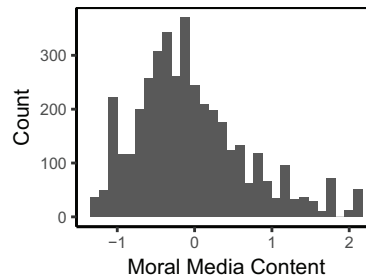


Figure B.5: Histogram of moralization in individual media environments. The variable is computed by averaging general MFT Scores for all media sources regularly consumed by each respondent.

### B.2.3 Remaining Variables in 2012 ANES

The 2012 ANES contains two representative cross-sectional samples which are pooled in the analyses. One sample was collected via computer assisted face-to-face interviews while the other is based on an internet panel. Most items described below are drawn from the pre-election wave of the survey.<sup>3</sup> The key independent variable used to predict the emphasis on each moral foundation in the first step of the analyses is *political ideology*. Respondents were asked to place themselves on a seven-point scale ranging from extremely liberal to extremely conservative, which was transformed into dichotomous indicators for respondents who identified as liberals, conservatives, or moderates. Additional control variables included in the analyses are *age*, *sex*, *race* (African American), *church attendance*, survey mode (online vs. offline), *education* (college degree), as well as the overall length of the individual responses in the open-ended questions (*measured as logged number of words*). Furthermore, the 2012 ANES included the *Wordsum* vocabulary test as a measure of literacy and verbal skill. It consists of a series of items asking respondents to choose a term that is closest to a target word. The *Wordsum* score consists of an additive index of correct responses in ten individual trials. The inclusion of education, the length of individual responses, and the *Wordsum* score as control variables should account for potential confounding factors such as general effects of increased political literacy on the complexity of open-ended responses.

In order to examine the political relevance of moral reasoning measured through open-ended responses, the MFT scores for each moral foundation are used as independent variables to predict political outcomes. The dependent variable considered in the main text is *voting behavior* (measured as a dichotomous indicator of vote choice for the Democratic rather than the Republican Presidential candidate reported in the post-election wave). Supplementary analyses in the appendix additionally examine *candidate* and *party evaluations*, each measured as the respective feeling thermometer differentials. In addition to the controls discussed previously, these analyses include measures of *party identification*, which were recoded similarly to ideology.

The last step of the analyses investigates how the expression of moral considerations in political judgment is influenced by the content of individual media environments. Additional control variables in this step include *political knowledge* (measured as the sum of correct answers to factual knowledge questions), *political media exposure* (measured as the sum of weekly news consumption through TV, radio, internet, and print), and the frequency of *political discussions* with friends and family members. As discussed in the main text, the analyses explore whether these factors influence *general* moral reasoning. Figure B.6 provides histograms of the variables included in each stage of the analyses. With the exception of age, all independent variables that were treated as continuous were rescaled to range from 0 to 1.

<sup>3</sup>The open-ended items were included only in the pre-election wave. Accordingly, wherever possible, the set of explanatory variables was limited to the pre-election wave.

## APPENDIX B. SUPPLEMENTARY MATERIAL FOR CHAPTER 3

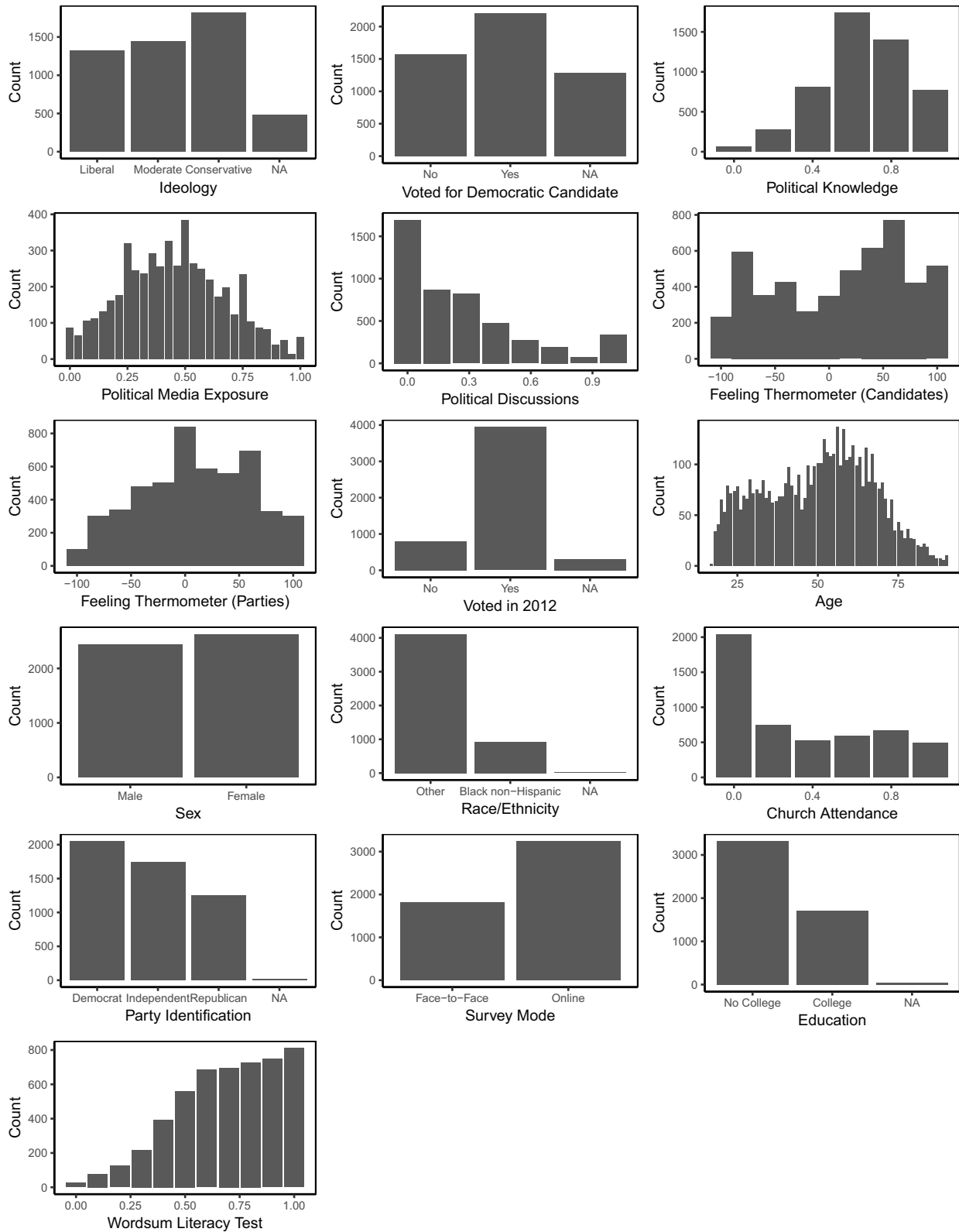


Figure B.6: Histograms of variables included in analyses.

### B.2.4 Open-ended Responses and MFT Scores in Replication Sample

The survey for the replication analysis was conducted via telephone with 594 adults aged 18 or older between early January, 2001 and July, 2003. The telephone numbers were a random-digit-dial (RDD) sample drawn from residents within a 25 mile radius of a large northeastern state university. As such, the survey was not conducted during a major presidential election campaign and under a Republican presidency. Furthermore, the survey varied the set of open-ended items. Rather than asking about attitudes towards presidential candidates and both major parties, respondents were asked to describe liberals and conservatives as well as their respective beliefs in general:

- “Can you briefly describe [liberals/conservatives] in your own words? What are they like?”
- “Can you briefly describe the political beliefs of [liberals/conservatives] in your own words? What do they believe?”

Based on these items, MFT scores were computed using the same procedures as for the 2012 ANES (i.e., pre-processing, weighting, etc.). Figure B.7 displays the proportions of individuals in the replication dataset who mentioned each moral foundation in their open-ended response. Compared to the ANES sample (see Figure B.2), fewer individuals mentioned any of the foundations, which can be explained by the fact that the average response length in the telephone study is substantially shorter (28 words) than in the ANES (75 words). Notwithstanding, each of the foundations (except sanctity), was mentioned by between 10 and 20% of the respondents.

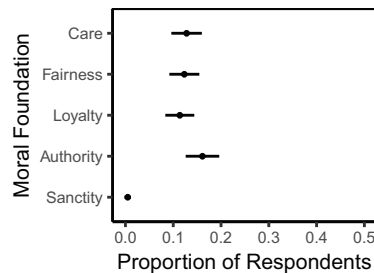


Figure B.7: Proportion of respondents mentioning each of the moral foundations in any of their open-ended responses, along with 95% confidence intervals in the replication dataset (RDD adult sample).

### B.2.5 Control Variables in Replication Sample

The coding of the remaining variables in the RDD survey is equivalent to those in the ANES analysis, although the survey did not contain the Wordsum scores (and varying survey mode) as additional controls. Histograms for all variables in the replication dataset are displayed in Figure B.8.

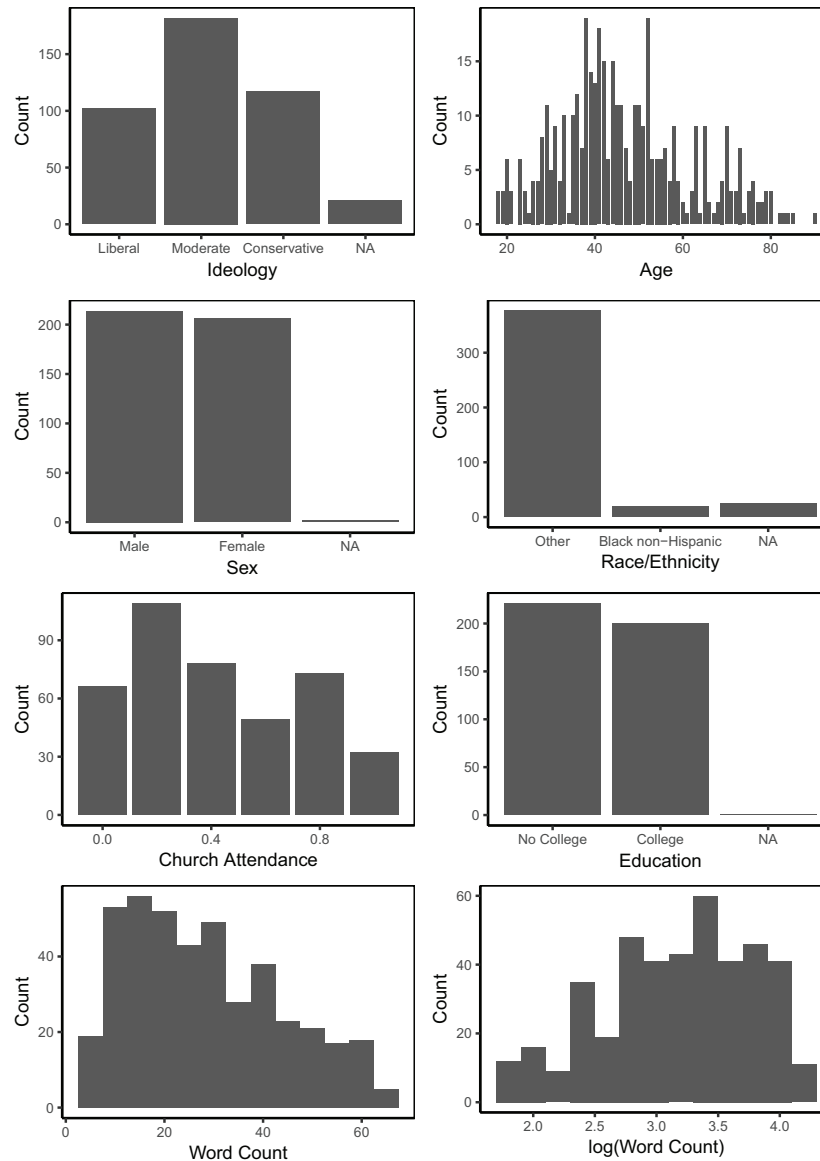


Figure B.8: Histograms of variables included in replication survey.

## B.3 Additional Model Results & Robustness Checks

### B.3.1 Replicating Ideological Differences using RDD Adult Sample

The article raised the possibility that terms in the dictionary may coincidentally recover unrelated differences in word choice between liberals and conservatives when discussing their attitudes towards parties and candidates in the 2012 U.S. Presidential election. For example, one prominent issue in the election was the Affordable Care Act, which might increase the likelihood of Democrats mentioning the term “care” and thereby increasing the emphasis on the harm/care foundation irrespective of underlying moral considerations. In that case, observed ideological differences may be an artifact of the context in which the survey took place.

To address this concern, I replicated the analysis from Figure 1 in the article using data from a separate survey conducted via telephone with 594 adults aged 18 or older between early January, 2001 and July, 2003. The telephone numbers were a random-digit-dial (RDD) sample drawn from residents within a 25 mile radius of a large northeastern state university.<sup>4</sup> The open-ended items asked respondents to describe liberals and conservatives as *social groups* as well as their respective *beliefs* in general. The coding and analyses are equivalent to those for Figure 1, although the survey did not contain the Wordsum scores included in the main analyses.

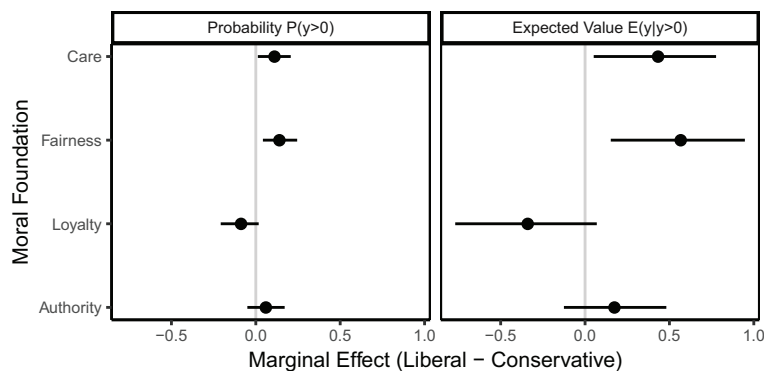


Figure B.9: Replication of main model (c.f., Figure 3.1) using RDD adult sample. Figure displays difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel), and in the MFT score given that the foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables include church attendance, education, age, sex, race, and response length. Full model results are displayed in the appendix, Table B.6.

Figure B.9 shows patterns that are consistent with previous results. Liberals are more likely to emphasize the foundations of care and fairness. The result for the loyalty dimension, however, do not reach conventional levels of statistical significance. Additional analyses reveal that the ideological differences in moral reasoning are mostly due to the fact that respondents who identify as liberals emphasize the foundations of care and fairness more strongly than conservatives when describing their ingroup (i.e., other liberals and their beliefs), while conservatives

<sup>4</sup>I thank Leonie Huddy for sharing these data.



emphasize the loyalty foundation more strongly than liberals when describing their ingroup (results available upon request). The fact that the same basic ideological pattern can be recovered in a survey that was conducted in a different political context (non-election period, Republican administration), employed a different survey mode (phone interview), and relied on a different set of open-ended survey questions (asking about liberals and conservatives and their respective beliefs), provides additional evidence that the MFT dictionary recovers basic moral considerations in political attitude expression.

### B.3.2 Negations and Valence in Open-Ended Responses (2012 ANES)

Dictionary-based approaches do not take into account the context in which signal terms appear. As such, they cannot capture directly whether individuals use certain words in a positive context or as part of a negation. Interpreting MFT scores as endorsements of a foundation might therefore be problematic if respondents commonly reject certain signal terms. Manual inspections of a random subset of open-ended suggested that most appearances of dictionary terms are not used in a direct negation (see also Table B.2). To provide a more comprehensive overview of potential negations or valenced statements in open-ended responses, I repeated the analyses from Figure 1 for different subsets of the questionnaire as well as the dictionary. The results are displayed in Figure B.10.

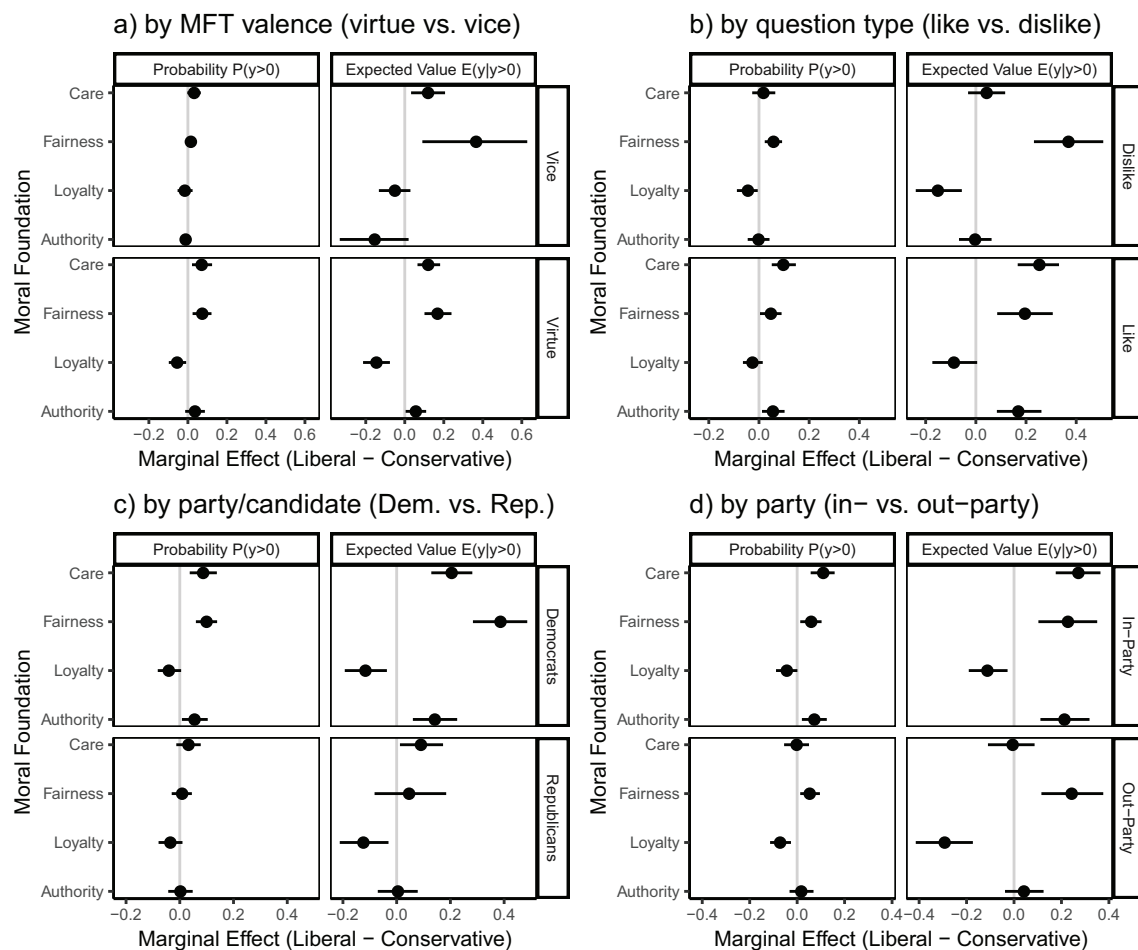


Figure B.10: Replication of main model (c.f., Figure 1) by different subgroups. The figure displays difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel), and in the MFT score given that the foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables include church attendance, education, age, sex, race, and response length. Full model results are displayed in Tables B.7 to B.14.

Panel a) divides the MFT dictionary in positive and negative terms for each moral foundation (i.e. virtues and vices, c.f., <http://moralfoundations.org>) before examining ideological differences. Whether we only consider signal terms that represent “vices” (top) or “virtues” (bottom), the differences between liberals and conservatives regarding individualizing foundations (care, fairness) are largely consistent. However, there is suggestive evidence that liberals are more likely to discuss authority in terms of virtues, while conservatives are more likely to emphasize the foundation in terms of vices. Panel b) uses the full MFT dictionary again, but now splits up the set of open-ended questions before repeating the analysis. The top part displays the ideological differences when focusing on “dislike” questions, whereas the bottom part only considers “like” questions. While the results are again largely consistent, some interesting differences appear. For example, liberals are only more likely to emphasize the care foundation than conservatives when discussing issues they *like* about parties and candidates, and not when they talk about things they *dislike*. It seems unlikely that the consistent differences on “Likes” are driven by negations rather than positive endorsements of moral foundations. On the other hand, however, we also observe that liberals are more likely to emphasize the authority foundation than conservatives when talking about aspects they like. While this finding contradicts MFT, the remaining two panels provide some intuition about possible explanations. Panel c) splits the open-ended responses by the party each respondent is asked to discuss. Here we observe that the ideological differences in moral reasoning are more pronounced when discussing the Democratic party and candidate as compared to the Republican party and candidate. A similar distinction is made in Panel d), where open-ended responses are divided into description of the respondent’s respective out-party vs. in-party. Now, ideological differences appear to be more pronounced when respondents discuss their in-party rather than their out-party. Overall, regarding the surprising finding on authority, it appears that liberals were more likely to emphasize the foundation when discussing aspects that they like about their in-party candidate. A plausible example for such a pattern could be that they were more likely to describe the Democratic presidential candidate (Barack Obama) as a good “leader”, which is a signal term for the authority dimension.

### B.3.3 MFT and Party/Candidate Evaluations (2012 ANES)

The second part of the analyses in the main text examines whether the expression of moral foundations in open-ended responses is related to voting behavior. To further corroborate these results, I additionally examine the relationship of moral reasoning and attitudes towards political parties and candidates. Figure B.11 presents OLS estimates where feeling thermometer differentials between the Republican and the Democratic party (left panel) and between both Presidential candidates (right panel) are regressed on MFT scores for all moral foundations.

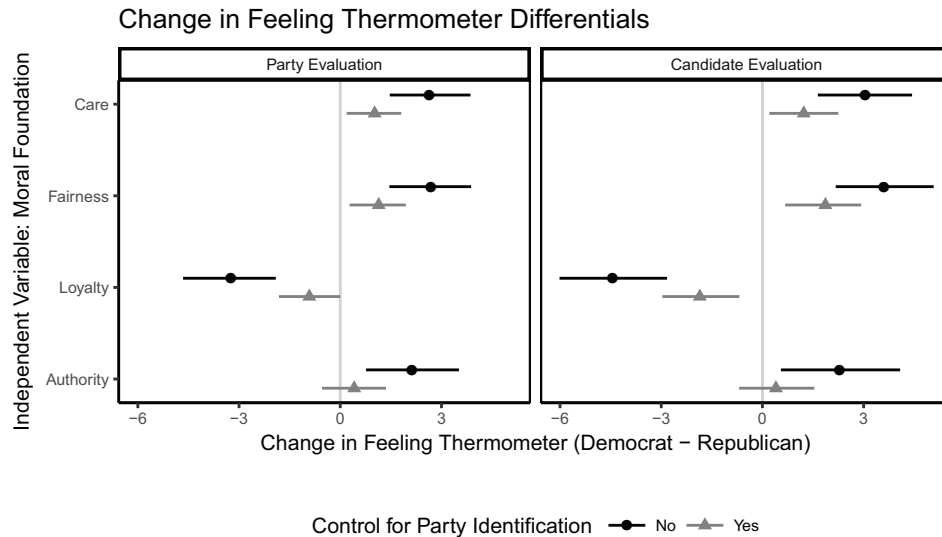


Figure B.11: Change in predicted feeling thermometer differential when MFT score is increased from its minimum (no overlap between dictionary and response) by one standard deviation, holding control variables constant at their respective means (along with 95% confidence intervals). Positive values indicate that respondents who emphasized the respective foundation evaluated the Democratic candidate/party more favorably than the Republican candidate/party, and vice versa. Estimates are based on a single OLS model (using robust standard errors) including MFT scores for each foundation and gray triangles indicate estimates while additionally controlling for party identification. Remaining control variables are age, sex, race, church attendance, survey mode, education, response length, and the Wordsum vocabulary score. Full model results are displayed in Table B.15.

Positive values indicate more favorable evaluations for the Democratic candidate or party and negative values indicate more favorable evaluations of the Republican candidate or party. The patterns are consistent with previous results. Individuals who emphasize considerations related to care and fairness evaluate the Democratic party/candidate on average about 3 points higher than the Republican party/candidate (on a 100 point scale). On the other hand, if individuals emphasized the loyalty dimension, they reported stronger preferences for the Republican party/candidate. Interestingly, mentioning terms that belong to the authority dimension appears to increase favorability towards the democratic party and candidate, which contradicts MFT. However, the effect disappears when party identification is included as a control variable.

### B.3.4 Other Correlates of General Moralization in Attitude Expression

The last analysis in the article focuses on the effect of moralization in individual media environments on general moral reasoning in open-ended responses. Additional controls included in the model were political knowledge, general media exposure, and discussion frequency. Figure B.12 allows for a comparison of effect sizes when each variable is increased from its empirical minimum value to its empirical maximum value, holding all other control variables constant at their means. To reiterate, the dependent variable captures the general tendency to emphasize *any* moral foundation. Estimates are based on a Tobit model and the estimated effects are decomposed into the probability of mentioning any moral foundation (left panel) as well as the emphasis on morality, given that any foundation was mentioned (right panel).

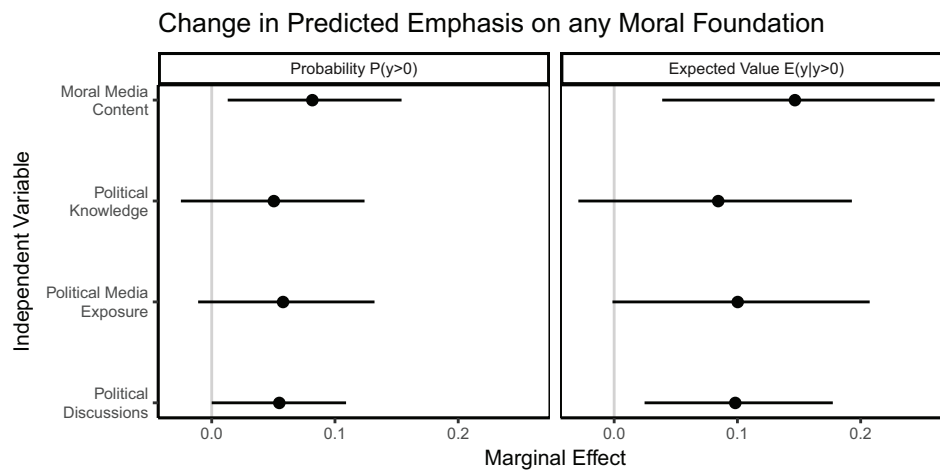


Figure B.12: Change in predicted overall reliance on moral foundations depending on moral media content, political knowledge, media exposure, and frequency of political discussions. The plot shows differences in predicted probabilities of mentioning any moral foundation (left panel) as well as in the summed MFT scores given that any foundation was mentioned (right panel), if each of the independent variables is increased from its minimum to its maximum value holding all other variables constant at their respective means (along with 95% confidence intervals). Additional control variables include age, sex, race, church attendance, survey mode, education, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix, Table B.5.

The significant positive effect of frequent political discussions (even after controlling for moral media content, political knowledge, and media exposure), is especially interesting. Citizens who engage in frequent political arguments are more likely to use moral considerations when evaluating candidates and parties, which could suggest that morality serves as a rhetorical tool utilized to convince others of certain political views.

### B.3.5 Comparing General Media MFT Scores with Manual Coding

The most fundamental concern related to the effects of individual media environments might be whether the content analysis of media sources using a dictionary is able to capture overall levels of moralization in news reporting. Luckily, a study reported in Feinberg and Willer (2013) included manual coding of a selection of newspaper articles on environmental issues to capture whether they use rhetoric grounded in each of the moral domains. Their coding therefore focuses on the same foundations without utilizing the dictionary. I computed a general moralization variable by summing the scores used in Feinberg and Willer (2013) and compared them to the MFT scores based on the procedures outlined above.<sup>5</sup>

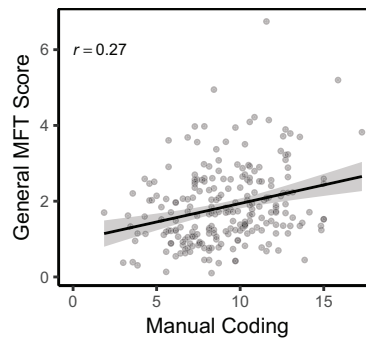


Figure B.13: Validity check comparing general MFT Scores with individual coder assessments of newspaper articles in Feinberg and Willer (2013).

Figure B.13 presents the correlation of general moralization in each article based on the manual coding in Feinberg and Willer (2013) compared to the dictionary method used in the analyses presented in the article. While the correlation is far from being perfect, the weighted dictionary method clearly captures some of the same variance as manual assessments of the emphasis on moral foundations. This is especially noteworthy since the coders in Feinberg and Willer (2013) did not rely on the moral foundations dictionary itself. This correspondence between subjective assessments and dictionary-based coding even persists when examining each moral foundation separately (c.f., Figure B.14).

<sup>5</sup>I thank Matthew Feinberg and Robb Willer for sharing these data.

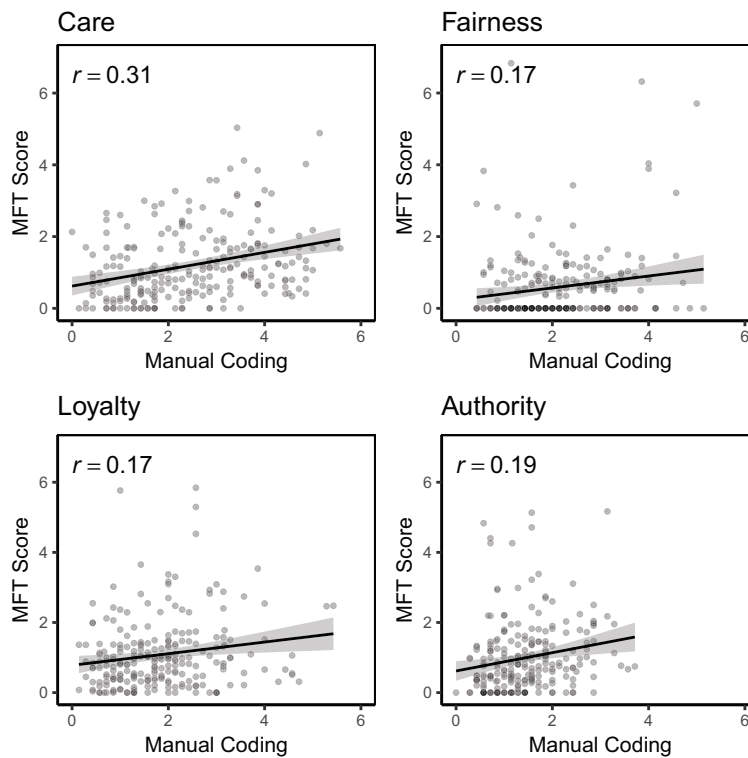


Figure B.14: Validity check comparing individual MFT Scores with individual coder assessments of newspaper articles in Feinberg and Willer (2013).

B.3.6 Face Validity of Sample Responses

Table B.2: Sample of open-ended responses in the 2012 ANES. Each row represents one individual and the columns display his or her raw verbatim response to each likes/dislikes item included in the analyses. Individuals in this table were selected if their total response length was within 10 words of the average (~ 75 words) and if they scored high on one of the moral foundations (see first column). Dictionary terms for the respective foundation are underlined>.

Foundation	Obama		Romney		Democratic Party		Republican Party	
	Like	Dislike	Like	Dislike	Like	Dislike	Like	Dislike
Care	supports ending war, supports affordable health care for all, supports the preservation of medicare and social security, looking into energy conservation to preserve our planet for future generations, initiatives to promote education and job growth and much more.	imposing a fine if someone does not get a health care plan	he supports a strong military	anti same sex marriage, anti women's choice for abortion, not supportive of health care reform act, not sensitive to the needs of the very poor and immigrants				
Fairness	people rights, economy, taxes for working people, understanding of international problems			abortion rights, women rights, tax breaks for the rich, military hawk, rude and condescending to President Obama	economy, women's rights, gay rights, health care, tax plan for working class, international strategy.			racist, elitist, trying to enrich the rich even more by hurting working people, international relations, health care, women rights, gay rights
Loyalty		He doesn't do enough about keeping us safe from our foreign enemies; he's too iffy about Israel; too favorable about homosexuality, abortion.	Like his close family ties; good ideas about keeping us safe from foreign countries; strong on Israel;				I like the people that are their leaders; I've never been disappointed in their positions on most things.	
Authority	competent, intelligent, but not strong in protecting US border, seriously dealing with illegal immigrant not rewarding for breaking the law. also, health care bill have me a little concern	lack of strong laws dealing with illegal aliens and US borders, strong and firm dealing with foreign countries in middle east, china, mexico; the health care--not too comfortable with what I am hearing about it.						



## B.4 Tables of Model Estimates

### B.4.1 Ideological Differences

Table B.3: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure 3.1 in the main text.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-0.391 (0.089)	-0.808 (0.145)	0.384 (0.117)	-0.131 (0.091)
Ideology (Moderate)	-0.225 (0.09)	-0.706 (0.149)	0.054 (0.123)	-0.050 (0.093)
Church Attendance	0.006 (0.103)	0.073 (0.169)	0.257 (0.134)	-0.110 (0.105)
Education (College Degree)	-0.108 (0.078)	0.289 (0.126)	0.342 (0.101)	0.109 (0.079)
Age	0.000 (0.002)	0.001 (0.004)	-0.007 (0.003)	0.002 (0.002)
Sex (Female)	0.093 (0.07)	0.118 (0.116)	-0.221 (0.093)	-0.105 (0.072)
Race (African American)	0.103 (0.101)	-0.042 (0.169)	-0.246 (0.138)	0.333 (0.102)
Word Count (log)	0.422 (0.04)	0.605 (0.067)	0.755 (0.055)	0.503 (0.042)
Wordsum Score	0.745 (0.186)	0.776 (0.309)	0.600 (0.247)	0.308 (0.19)
Survey Mode (Online)	-0.033 (0.084)	0.306 (0.14)	0.143 (0.111)	0.294 (0.087)
Intercept	-2.669 (0.226)	-5.055 (0.388)	-4.748 (0.313)	-3.231 (0.236)
log(Sigma)	0.671 (0.02)	1.045 (0.026)	0.885 (0.022)	0.684 (0.02)
N	4489	4489	4489	4489
Log-Likelihood	-5149	-3984	-4580	-5030

## B.4.2 Moral Considerations and Vote Choice

Table B.4: Logit models predicting democratic vote choice based on MFT score for each foundation. Positive coefficients indicate higher likelihood to vote for the Democratic candidate than the Republican candidate. Standard errors in parentheses. Estimates are used for Figure 3.2 in the main text.

Variable	(1)	(2)
Harm	0.192 (0.043)	0.171 (0.061)
Fairness	0.170 (0.041)	0.129 (0.053)
Ingroup	-0.190 (0.041)	-0.068 (0.054)
Authority	0.076 (0.041)	0.025 (0.056)
PID (Democrat)		2.618 (0.136)
PID (Republican)		-2.676 (0.156)
Church Attendance	-1.614 (0.113)	-1.352 (0.158)
Education (College Degree)	0.155 (0.085)	0.370 (0.119)
Age	-0.009 (0.002)	-0.017 (0.003)
Sex (Female)	0.275 (0.078)	0.160 (0.108)
Race (African American)	4.229 (0.263)	3.234 (0.288)
Word Count (log)	0.145 (0.044)	0.116 (0.061)
Wordsum Score	0.048 (0.21)	0.174 (0.29)
Survey Mode (Online)	-0.366 (0.095)	-0.388 (0.131)
Intercept	0.313 (0.238)	0.509 (0.328)
N	3706	3698
Log-Likelihood	-1955	-1138

### B.4.3 Media Content and Exposure to Moral Rhetoric

Table B.5: Tobit model predicting overall reliance on moral foundations (sum of MFT scores) based on media moralization, political knowledge, media exposure, and frequency of political discussions. Positive coefficients indicate stronger emphasis on any foundation. Standard errors in parentheses. Estimates are used for Figure 3.3 in the main text as well as Figure B.12 in the appendix.

Variable	(1)
Moral Media Content	0.086 (0.032)
Political Knowledge	0.180 (0.118)
Political Media Exposure	0.204 (0.11)
Political Discussions	0.195 (0.076)
Church Attendance	0.062 (0.062)
Education (College Degree)	0.088 (0.05)
Age	-0.001 (0.001)
Sex (Female)	0.014 (0.045)
Race (African American)	0.087 (0.058)
Word Count (log)	0.030 (0.025)
Wordsum Score	0.266 (0.119)
Survey Mode (Online)	0.068 (0.052)
Intercept	-0.014 (0.136)
log(Sigma)	0.210 (0.015)
N	3648
Log-Likelihood	-5240

## B.4.4 Additional Model Results &amp; Robustness Checks

## Replicating Ideological Differences using RDD Adult Sample

Table B.6: Tobit models predicting MFT score for each foundation based on ideology (telephone survey replication). Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.9 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-2.337 (1.052)	-3.342 (1.184)	1.703 (1.071)	-0.858 (0.786)
Ideology (Moderate)	-1.329 (0.839)	-2.280 (0.897)	-1.364 (1.046)	-1.075 (0.681)
Church Attendance	-0.887 (1.224)	0.317 (1.289)	0.831 (1.311)	1.095 (0.952)
Education (College Degree)	0.783 (0.753)	1.113 (0.805)	1.138 (0.842)	1.137 (0.597)
Age	-0.002 (0.025)	0.018 (0.026)	-0.029 (0.028)	-0.049 (0.021)
Sex (Female)	-0.761 (0.742)	0.748 (0.792)	-0.718 (0.816)	-0.677 (0.582)
Race (African American)	0.526 (1.648)	1.484 (1.587)	-0.061 (1.988)	0.570 (1.253)
Word Count (log)	2.182 (0.685)	1.113 (0.67)	1.058 (0.695)	0.653 (0.48)
Intercept	-10.402 (2.785)	-9.219 (2.828)	-8.164 (2.869)	-3.550 (1.893)
log(Sigma)	1.460 (0.121)	1.453 (0.137)	1.544 (0.126)	1.286 (0.109)
N	366	366	366	366
Log-Likelihood	-222	-187	-216	-254

Negations and Valence in Open-Ended Responses (2012 ANES)

a) by MFT valence (*virtue vs. vice*)

Table B.7: Virtues only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-0.443 (0.106)	-0.745 (0.153)	0.709 (0.178)	-0.196 (0.095)
Ideology (Moderate)	-0.342 (0.108)	-0.675 (0.157)	0.374 (0.187)	-0.053 (0.097)
Church Attendance	-0.024 (0.123)	0.029 (0.179)	0.589 (0.2)	-0.074 (0.11)
Education (College Degree)	-0.090 (0.093)	0.343 (0.133)	0.039 (0.152)	0.147 (0.083)
Age	0.001 (0.003)	0.000 (0.004)	-0.002 (0.004)	0.003 (0.002)
Sex (Female)	0.196 (0.084)	0.169 (0.122)	-0.129 (0.139)	-0.113 (0.075)
Race (African American)	0.102 (0.121)	-0.093 (0.179)	0.049 (0.202)	0.389 (0.106)
Word Count (log)	0.406 (0.048)	0.631 (0.071)	1.287 (0.087)	0.520 (0.044)
Wordsum Score	1.030 (0.224)	0.684 (0.326)	0.149 (0.369)	0.314 (0.198)
Survey Mode (Online)	-0.115 (0.1)	0.372 (0.148)	0.242 (0.166)	0.277 (0.091)
Intercept	-3.291 (0.274)	-5.354 (0.412)	-8.851 (0.517)	-3.451 (0.248)
log(Sigma)	0.803 (0.022)	1.084 (0.027)	1.171 (0.029)	0.717 (0.021)
N	4489	4489	4489	4489
Log-Likelihood	-4682	-3830	-3505	-4914

Table B.8: Vices only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-0.674 (0.244)	-3.120 (1.191)	0.276 (0.22)	1.211 (0.719)
Ideology (Moderate)	-0.058 (0.245)	-2.301 (1.224)	-0.283 (0.237)	-0.166 (0.783)
Church Attendance	0.037 (0.283)	1.000 (1.392)	-0.240 (0.26)	-1.755 (0.859)
Education (College Degree)	-0.266 (0.213)	0.265 (1.031)	1.024 (0.191)	-0.672 (0.63)
Age	-0.003 (0.006)	0.038 (0.03)	-0.017 (0.006)	-0.007 (0.018)
Sex (Female)	-0.201 (0.193)	-1.248 (0.961)	-0.422 (0.178)	-0.415 (0.578)
Race (African American)	0.196 (0.272)	0.506 (1.372)	-0.893 (0.287)	0.129 (0.864)
Word Count (log)	1.635 (0.126)	2.798 (0.602)	0.966 (0.106)	3.162 (0.398)
Wordsum Score	-0.051 (0.509)	4.297 (2.65)	1.496 (0.488)	1.091 (1.571)
Survey Mode (Online)	0.226 (0.228)	-0.923 (1.114)	0.059 (0.213)	1.320 (0.705)
Intercept	-10.814 (0.735)	-37.200 (4.374)	-8.292 (0.634)	-28.379 (2.687)
log(Sigma)	1.381 (0.035)	2.483 (0.08)	1.343 (0.033)	2.108 (0.065)
N	4489	4489	4489	4489
Log-Likelihood	-2672	-867	-2993	-1168

## APPENDIX B. SUPPLEMENTARY MATERIAL FOR CHAPTER 3

b) *by question type (like vs. dislike)*

Table B.9: Likes only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-1.108 (0.173)	-0.982 (0.282)	0.456 (0.232)	-0.756 (0.204)
Ideology (Moderate)	-0.780 (0.175)	-1.036 (0.293)	0.039 (0.246)	-0.350 (0.207)
Church Attendance	-0.176 (0.204)	-0.243 (0.333)	0.567 (0.267)	0.286 (0.235)
Education (College Degree)	0.064 (0.151)	0.636 (0.246)	0.715 (0.2)	0.280 (0.177)
Age	0.000 (0.004)	0.011 (0.007)	-0.008 (0.006)	0.009 (0.005)
Sex (Female)	0.024 (0.138)	0.186 (0.227)	-0.069 (0.185)	-0.265 (0.162)
Race (African American)	0.418 (0.194)	0.028 (0.332)	0.183 (0.266)	0.403 (0.229)
Word Count (log)	0.810 (0.081)	1.030 (0.132)	1.502 (0.115)	0.957 (0.095)
Wordsum Score	0.951 (0.367)	1.160 (0.609)	0.458 (0.493)	0.617 (0.427)
Survey Mode (Online)	-0.244 (0.163)	0.597 (0.276)	-0.181 (0.217)	0.439 (0.195)
Intercept	-5.737 (0.461)	-10.578 (0.796)	-10.582 (0.678)	-7.830 (0.553)
log(Sigma)	1.214 (0.026)	1.626 (0.031)	1.419 (0.03)	1.371 (0.027)
N	4489	4489	4489	4489
Log-Likelihood	-4097	-3512	-3445	-4299

Table B.10: Dislikes only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-0.193 (0.162)	-2.136 (0.428)	0.739 (0.229)	0.015 (0.146)
Ideology (Moderate)	-0.036 (0.166)	-1.458 (0.433)	0.197 (0.244)	0.151 (0.149)
Church Attendance	0.163 (0.187)	0.188 (0.505)	0.142 (0.262)	-0.428 (0.17)
Education (College Degree)	-0.278 (0.142)	0.308 (0.37)	0.497 (0.196)	0.158 (0.126)
Age	0.000 (0.004)	-0.037 (0.011)	-0.011 (0.006)	0.003 (0.004)
Sex (Female)	0.201 (0.128)	0.237 (0.341)	-0.494 (0.181)	-0.087 (0.116)
Race (African American)	0.041 (0.186)	-1.070 (0.52)	-0.977 (0.287)	0.570 (0.163)
Word Count (log)	0.832 (0.075)	1.978 (0.21)	1.247 (0.109)	0.971 (0.07)
Wordsum Score	1.115 (0.342)	1.994 (0.926)	1.198 (0.489)	0.391 (0.306)
Survey Mode (Online)	0.109 (0.153)	0.327 (0.406)	0.675 (0.222)	0.380 (0.14)
Intercept	-6.452 (0.438)	-15.081 (1.235)	-9.812 (0.65)	-6.684 (0.407)
log(Sigma)	1.165 (0.025)	1.915 (0.038)	1.428 (0.029)	1.050 (0.025)
N	4489	4489	4489	4489
Log-Likelihood	-4378	-2777	-3738	-4065



## APPENDIX B. SUPPLEMENTARY MATERIAL FOR CHAPTER 3

c) *by party/candidate (Democrat vs. Republican)*

Table B.11: Democratic party/candidate only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-0.820 (0.152)	-1.998 (0.267)	0.545 (0.191)	-0.607 (0.172)
Ideology (Moderate)	-0.510 (0.154)	-0.945 (0.26)	0.031 (0.203)	-0.368 (0.176)
Church Attendance	0.240 (0.176)	0.355 (0.306)	0.457 (0.219)	0.191 (0.199)
Education (College Degree)	-0.147 (0.133)	0.472 (0.226)	0.505 (0.164)	0.078 (0.15)
Age	-0.003 (0.004)	-0.004 (0.006)	-0.009 (0.005)	0.009 (0.004)
Sex (Female)	0.109 (0.121)	0.113 (0.208)	-0.247 (0.152)	-0.315 (0.137)
Race (African American)	-0.087 (0.174)	0.079 (0.299)	-0.261 (0.225)	0.339 (0.194)
Word Count (log)	0.644 (0.07)	1.081 (0.124)	1.169 (0.092)	0.948 (0.081)
Wordsum Score	0.986 (0.32)	1.304 (0.561)	0.401 (0.403)	0.599 (0.362)
Survey Mode (Online)	-0.163 (0.143)	0.771 (0.255)	0.065 (0.18)	0.591 (0.167)
Intercept	-4.548 (0.393)	-9.723 (0.735)	-8.063 (0.534)	-7.174 (0.471)
log(Sigma)	1.133 (0.024)	1.513 (0.033)	1.273 (0.028)	1.225 (0.025)
N	4489	4489	4489	4489
Log-Likelihood	-4698	-3169	-3781	-4437

Table B.12: Republican Party/Candidate only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-0.430 (0.189)	-0.265 (0.353)	0.638 (0.241)	-0.025 (0.175)
Ideology (Moderate)	-0.279 (0.193)	-1.331 (0.384)	0.188 (0.256)	0.209 (0.178)
Church Attendance	-0.317 (0.221)	-0.510 (0.423)	0.154 (0.277)	-0.352 (0.203)
Education (College Degree)	-0.208 (0.166)	0.467 (0.312)	0.587 (0.206)	0.377 (0.151)
Age	0.004 (0.005)	-0.004 (0.009)	-0.014 (0.006)	0.001 (0.004)
Sex (Female)	0.092 (0.151)	0.346 (0.289)	-0.380 (0.191)	0.055 (0.138)
Race (African American)	0.491 (0.213)	-1.009 (0.451)	-0.615 (0.296)	0.659 (0.194)
Word Count (log)	1.110 (0.091)	1.406 (0.171)	1.462 (0.118)	1.020 (0.083)
Wordsum Score	0.982 (0.402)	1.467 (0.779)	1.382 (0.52)	0.316 (0.366)
Survey Mode (Online)	0.156 (0.18)	0.081 (0.344)	0.466 (0.231)	0.262 (0.166)
Intercept	-8.220 (0.535)	-13.016 (1.03)	-11.071 (0.711)	-7.459 (0.485)
log(Sigma)	1.269 (0.028)	1.802 (0.035)	1.444 (0.031)	1.205 (0.027)
N	4489	4489	4489	4489
Log-Likelihood	-3818	-3075	-3372	-3957

## APPENDIX B. SUPPLEMENTARY MATERIAL FOR CHAPTER 3

d) *by party (in- vs. out-party)*

Table B.13: In-party only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	-1.136 (0.194)	-1.130 (0.31)	0.549 (0.208)	-0.924 (0.228)
Ideology (Moderate)	-0.792 (0.212)	-0.873 (0.348)	0.082 (0.241)	-0.222 (0.248)
Church Attendance	-0.355 (0.234)	-0.372 (0.376)	0.083 (0.247)	0.056 (0.272)
Education (College Degree)	-0.120 (0.178)	0.545 (0.283)	0.360 (0.189)	0.164 (0.208)
Age	0.005 (0.005)	0.006 (0.008)	-0.003 (0.005)	0.004 (0.006)
Sex (Female)	0.237 (0.16)	0.137 (0.257)	-0.055 (0.172)	-0.100 (0.187)
Race (African American)	0.159 (0.215)	-0.147 (0.358)	0.048 (0.238)	0.367 (0.251)
Word Count (log)	0.760 (0.094)	1.000 (0.152)	1.300 (0.109)	0.937 (0.111)
Wordsum Score	0.759 (0.423)	1.193 (0.691)	0.576 (0.461)	0.697 (0.496)
Survey Mode (Online)	-0.213 (0.19)	0.466 (0.313)	0.086 (0.205)	0.610 (0.229)
Intercept	-5.251 (0.54)	-9.540 (0.914)	-8.853 (0.652)	-7.264 (0.648)
log(Sigma)	1.173 (0.031)	1.548 (0.038)	1.163 (0.035)	1.331 (0.031)
N	2950	2950	2950	2950
Log-Likelihood	-2880	-2298	-2263	-3039

Table B.14: Out-party only: Tobit models predicting MFT score for each foundation based on ideology. Positive coefficients indicate stronger emphasis on the respective foundation. Standard errors in parentheses. Estimates are used for Figure B.10 in the appendix.

Variable	Care	Fairness	Loyalty	Authority
Ideology (Conservative)	0.026 (0.211)	-1.382 (0.368)	1.485 (0.329)	-0.184 (0.191)
Ideology (Moderate)	0.058 (0.236)	-1.359 (0.426)	0.631 (0.383)	0.076 (0.211)
Church Attendance	0.154 (0.251)	0.094 (0.449)	-0.016 (0.385)	-0.226 (0.229)
Education (College Degree)	-0.104 (0.193)	0.424 (0.336)	0.539 (0.292)	0.298 (0.174)
Age	-0.004 (0.005)	-0.019 (0.009)	-0.023 (0.008)	0.000 (0.005)
Sex (Female)	0.177 (0.174)	0.627 (0.309)	-0.682 (0.268)	-0.074 (0.158)
Race (African American)	0.079 (0.238)	-1.056 (0.447)	-1.076 (0.398)	0.677 (0.209)
Word Count (log)	0.913 (0.103)	1.615 (0.193)	1.354 (0.161)	0.965 (0.096)
Wordsum Score	1.257 (0.465)	0.324 (0.822)	1.251 (0.721)	0.620 (0.419)
Survey Mode (Online)	0.046 (0.209)	0.412 (0.366)	0.371 (0.325)	0.173 (0.189)
Intercept	-7.057 (0.609)	-11.600 (1.138)	-10.597 (0.965)	-6.935 (0.564)
log(Sigma)	1.251 (0.032)	1.609 (0.046)	1.588 (0.037)	1.130 (0.033)
N	2950	2950	2950	2950
Log-Likelihood	-2890	-1738	-2412	-2582

MFT and Party/Candidate Evaluations (2012 ANES)

Table B.15: OLS models predicting feeling thermometer differentials based on MFT score for each foundation. Positive coefficients indicate more favorable evaluation of Democratic candidate/party than the Republican candidate/party, and vice versa. Standard errors in parentheses. Estimates are used for Figure B.11 in the appendix.

Variable	Party (1)	Party (2)	Cand. (1)	Cand. (2)
Harm	2.661 (0.649)	0.997 (0.453)	3.056 (0.778)	1.249 (0.582)
Fairness	2.661 (0.647)	1.127 (0.451)	3.612 (0.777)	1.850 (0.58)
Ingroup	-3.244 (0.65)	-0.930 (0.453)	-4.444 (0.782)	-1.856 (0.585)
Authority	2.119 (0.669)	0.435 (0.466)	2.246 (0.796)	0.393 (0.595)
PID (Democrat)		45.081 (1.099)		48.219 (1.407)
PID (Republican)		-44.980 (1.22)		-52.268 (1.566)
Church Attendance	-27.723 (1.866)	-11.363 (1.324)	-35.802 (2.236)	-17.406 (1.702)
Education (College Degree)	-0.076 (1.492)	1.252 (1.04)	1.161 (1.791)	2.686 (1.338)
Age	-0.119 (0.041)	-0.124 (0.028)	-0.315 (0.049)	-0.317 (0.036)
Sex (Female)	7.447 (1.311)	2.843 (0.917)	9.269 (1.571)	4.238 (1.177)
Race (African American)	53.372 (1.784)	20.987 (1.325)	63.622 (2.135)	28.122 (1.698)
Word Count (log)	2.147 (0.736)	0.986 (0.513)	2.324 (0.882)	0.934 (0.659)
Wordsum Score	0.462 (3.412)	3.573 (2.382)	0.670 (4.095)	3.961 (3.066)
Survey Mode (Online)	-5.993 (1.544)	-2.185 (1.079)	-8.463 (1.849)	-4.310 (1.386)
Intercept	8.004 (3.762)	4.326 (2.671)	18.373 (4.505)	15.856 (3.426)
N	4924	4912	4938	4927
R-squared (adj.)	0.214	0.621	0.228	0.571

Appendix **C**

**Supplementary Material for Chapter 4**

### C.1 Basic Rules on the Subreddit ChangeMyView

Below is a summary of the set of rules to participate in discussions on /r/ChangeMyView as described in April 2018. The current rules can be viewed at <https://www.reddit.com/r/changemyview/wiki/rules>

#### Rules for submission of new discussion posts:

- A Explain the reasoning behind your view, not just what that view is (500+ characters required).
- B You must personally hold the view and demonstrate that you are open to it changing.
- C Submission titles must adequately sum up your view and include "CMV:" at the beginning.
- D Posts cannot express a neutral stance, carry a risk of personal endangerment, be self-promotional, or discuss this subreddit (visit r/ideasformv instead).
- E Only post if you are willing to have a conversation with those who reply to you, and are available to start doing so within 3 hours of posting.

#### Rules for commenting in existing discussions:

- 1 Direct responses to a CMV post must challenge at least one aspect of OPs stated view (however minor), or ask a clarifying question.
- 2 Don't be rude or hostile to other users.
- 3 Refrain from accusing OP or anyone else of being unwilling to change their view.
- 4 Award a delta if you've acknowledged a change in your view. Do not use deltas for any other purpose.
- 5 Comments must contribute meaningfully to the conversation.

## C.2 Moral Foundations Dictionary

### Sources:

Graham, Jesse, Jonathan Haidt, and Brian A. Nosek. 2009. "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology* 96 (5): 1029–1046, as well as <http://www.moralfoundations.org/>

### Note:

Terms with (\*) indicate that the word stem rather than the exact word was matched in the open-ended survey responses.

### Care

amity, benefit\*, care, caring, compassion\*, defen\*, empath\*, guard\*, peace\*, preserve, protect\*, safe\*, secur\*, shelter, shield, sympath\*, abandon\*, abuse\*, annihilate\*, attack\*, brutal\*, cruel\*, crush\*, damag\*, destroy, detriment\*, endanger\*, exploit, exploited, exploiting, exploits, fight\*, harm\*, hurt\*, impair, kill, killed, killer\*, killing, kills, ravage, ruin\*, spurn, stomp, suffer\*, violen\*, war, warl\*, warring, wars, wound\*

### Fairness

balance\*, constant, egalitar\*, equable, equal\*, equity, equivalent, evenness, fair, fair-\*, fairly, fairmind\*, fairness, fairplay, homologous, honest\*, impartial\*, justice, justifi\*, justness, reasonable, reciproc\*, rights, tolerant, unbiass\*, unprejudice\*, bias\*, bigot\*, discriminat\*, dishonest, disproportion\*, dissociate, exclud\*, exclusion, favoritism, inequitable, injust\*, preference, prejud\*, segregat\*, unequal\*, unfair\*, unjust\*, unscrupulous

### Loyalty

ally, cadre, cliqu\*, cohort, collectiv\*, communal, commune\*, communis\*, communit\*, comrad\*, devot\*, familial, families, family, fellow\*, group, guild, homeland\*, insider, joint, loyal\*, member, nation\*, patriot\*, segregat\*, solidarity, together, unison, unite\*, abandon\*, apostasy, apostate, betray\*, deceiv\*, deserted, deserter\*, deserting, disloyal\*, enem\*, foreign\*, immigra\*, imposter, individual\*, jilt\*, miscreant, renegade, sequester, spy, terroris\*, traitor\*, treacher\*, treason\*

### Authority

abide, allegian\*, authorit\*, bourgeoisie, caste\*, class, command, complian\*, comply, control, defer, defere\*, duti\*, duty, father\*, hierarch\*, honor\*, law, lawful\*, leader\*, legal\*, loyal\*, mother, mothering, motherl\*, mothers, obedien\*, obey\*, order\*, permission, permit, position, preserve, rank\*, respect, respected, respectful\*, respects, revere\*, serve, status\*, submi\*, supremacy, tradition\*, venerat\*, agitat\*, alienate, apostasy, apostate, betray\*, defector, defian\*, defy\*, denounce, deserted, deserter\*, deserting, disloyal\*, disobe\*, disrespect\*, dissent\*, dissident, heretic\*, illegal\*, insubordinat\*, insurgent, lawless\*, mutinous, nonconformist, obstruct, oppose, protest, rebel\*, refuse, remonstrate, riot\*, sediti\*, subver\*, traitor\*, treacher\*, treason\*, unfaithful



### Sanctity

abstemiousness, abstinence, abstinen\*, austerity, celiba\*, chast\*, church\*, clean\*, decen\*, holiness, holy, immaculate, innocent, integrity, limpid, maiden, modesty, piety, pious, preserve, pristine, pure\*, purity, refined, sacred\*, saint\*, steril\*, unadulterated, upright, virgin, virginal, virginity, virgins, virtuous, wholesome\*, adulter\*, apostasy, apostate, blemish, contagio\*, debase\*, debauché\*, defile\*, deprav\*, desecrat\*, dirt\*, disease\*, disgust\*, exploit, exploitat\*, exploited, exploiting, exploits, filth\*, gross, heretic\*, impiety, impious, indecen\*, intemperate, lax, lewd\*, obscen\*, pervert, profan\*, profligate, promiscu\*, prostitut\*, repuls\*, ruin\*, sick\*, sin, sinful\*, sinned, sinner\*, sinning, sins, slut\*, stain\*, taint\*, tarnish\*, tramp, trashy, unchaste, unclean\*, wanton, whore, wicked\*, wretched\*

### General Morality

bad, blameless, canon, character, commendable, correct, decen\*, doctrine, ethic\*, evil, exemplary, good, goodness, honest\*, ideal\*, immoral\*, indecen\*, integrity, laudable, lawful\*, legal\*, lesson, moral\*, noble, offend\*, offensive\*, piety, pious, praiseworthy, principle\*, proper, righteous\*, transgress\*, upright, upstanding, value\*, wholesome\*, wicked\*, worth\*, wretched\*, wrong\*

### C.3 Distribution of Moral Foundation Proportions in Paired Data

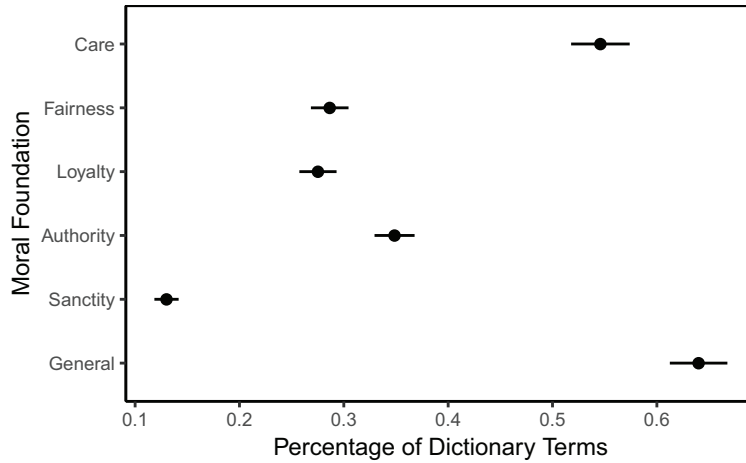


Figure C.1: Moral Foundations in Paired Data: Average percentage of dictionary terms relative to the total number of words in each original post starting a discussion (including 95% confidence intervals). Compared to the figure in the main text, this plot only includes opening statements that are part of the matched pair selection to analyze persuasive arguments.

## C.4 Structural Topic Model Results

### C.4.1 Original Posts

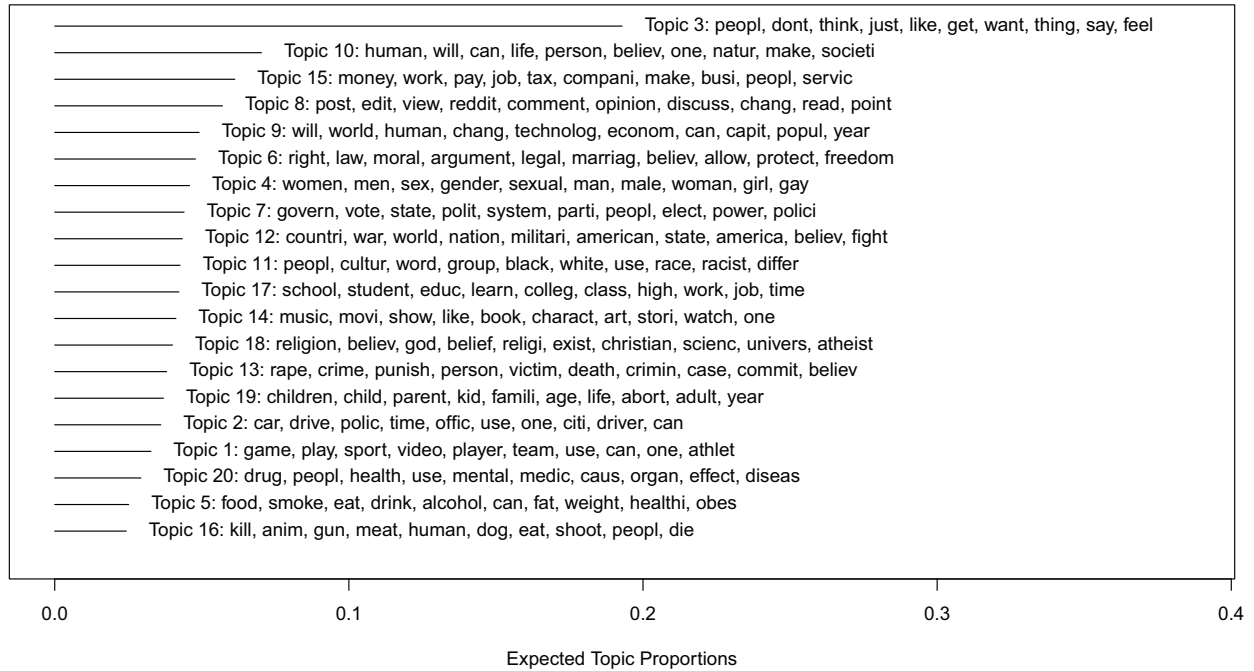


Figure C.2: Average topic proportions in opening statements on /r/ChangeMyView/ based on a structural topic model with 20 topics (c.f., Roberts et al., 2014). The plot additionally displays the ten most likely terms associated with each respective topic.

## C.4. STRUCTURAL TOPIC MODEL RESULTS

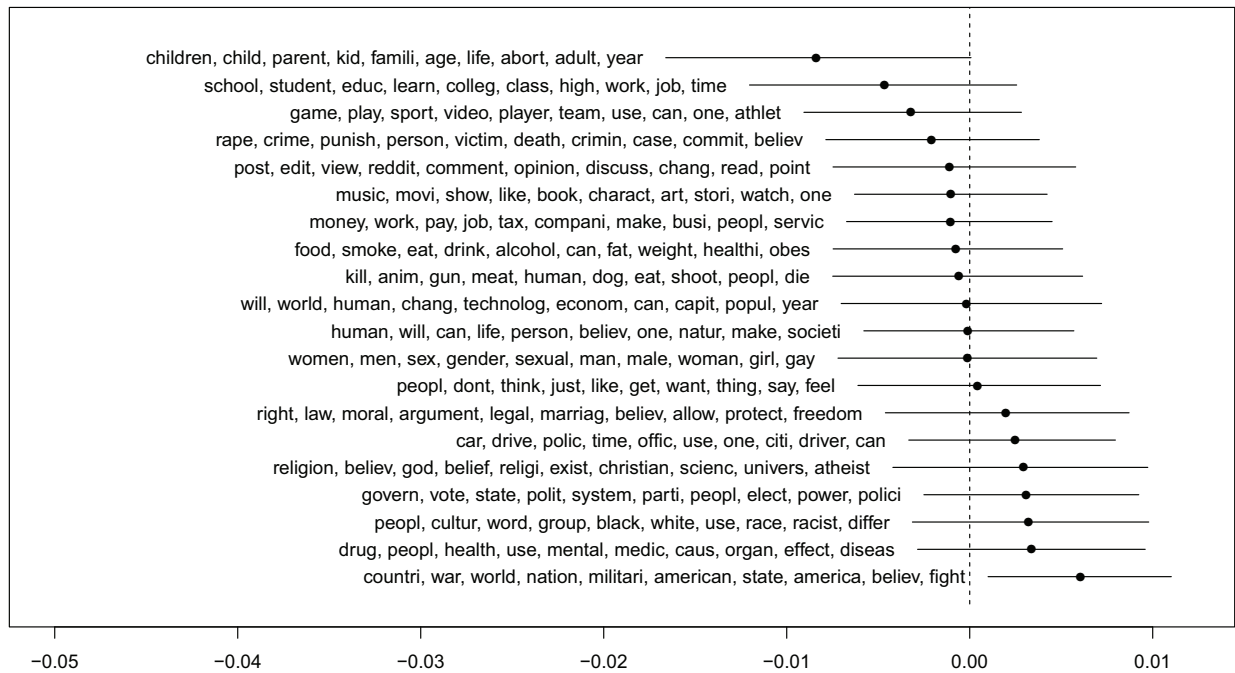


Figure C.3: Differences in topic proportions between opening statements on /r/ChangeMyView/ that resulted in opinion change ( $\Delta$  awarded) versus not (including 95% confidence intervals). Estimates are based on the structural topic model described in the previous figure. Positive values indicate higher topic prevalence among discussions that resulted in opinion change and vice versa. Labels are based on the ten highest probability terms related to the topic.

C.4.2 Responses Challenging the OP

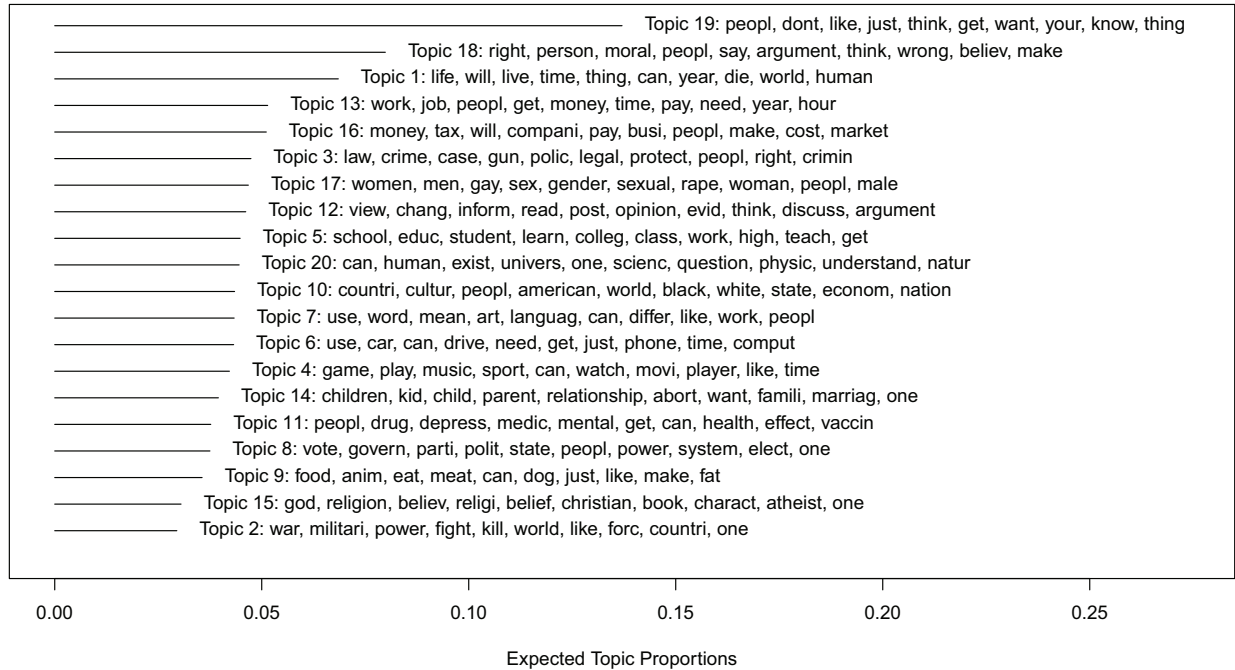


Figure C.4: Average topic proportions in posts challenging the OP on /r/ChangeMyView/ based on a structural topic model with 20 topics (c.f., Roberts et al., 2014). The plot additionally displays the ten most likely terms associated with each respective topic.

## C.4. STRUCTURAL TOPIC MODEL RESULTS

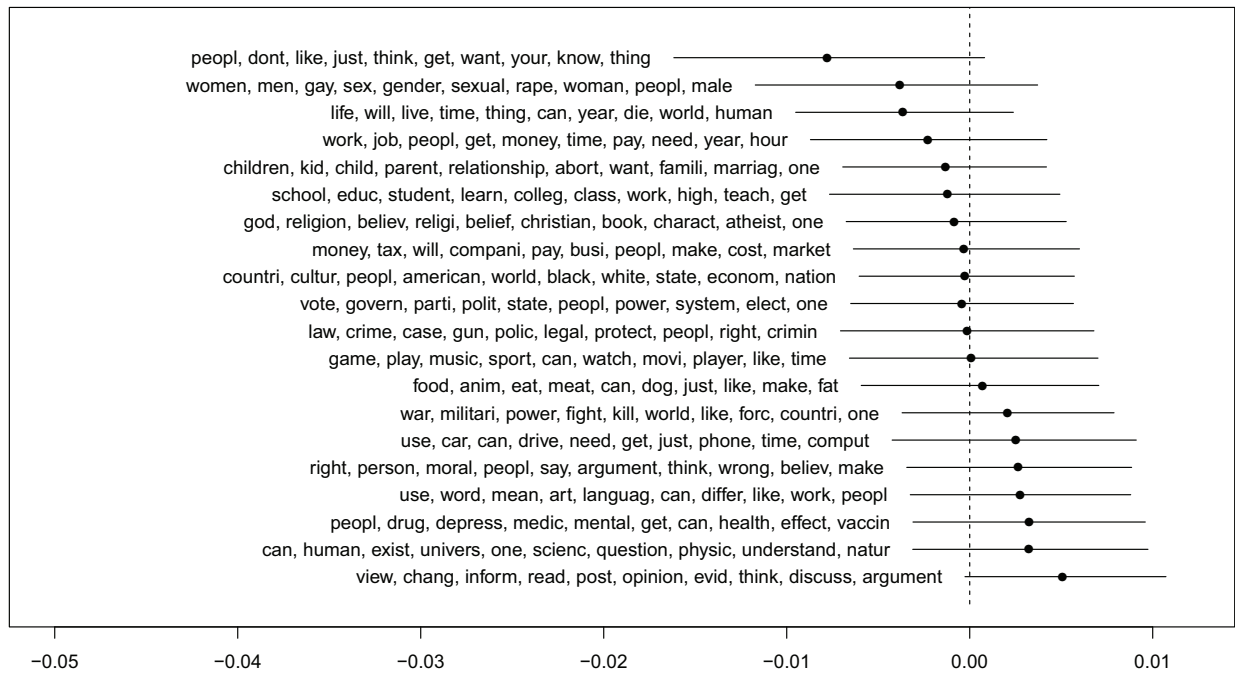


Figure C.5: Differences in topic proportions between persuasive and non-persuasive responses challenging the OP (including 95% confidence intervals). Estimates are based on the structural topic model described in the previous figure. Positive values indicate higher topic prevalence among posts that received a  $\Delta$  by the OP and vice versa. Labels are based on the ten highest probability terms related to the topic.

## C.5 Tables of Model Estimates

Table C.1: Logit models predicting argument persuasiveness as a function of moral word use (measured via MFT dictionary proportions). Positive coefficients indicate higher probability of changing the OPs' mind ( $\Delta$  awarded). Standard errors (clustered by discussion thread) in parentheses. Estimates are used for Figure 4.6 in the main text.

Variable	Full Response Path	Root Response	Truncated Root Response
Care	-0.004 (0.025)	-0.009 (0.023)	-0.007 (0.023)
Fairness	-0.029 (0.036)	-0.024 (0.033)	-0.032 (0.031)
Loyalty	0.005 (0.035)	0.017 (0.033)	0.018 (0.03)
Authority	0.003 (0.03)	-0.005 (0.028)	0.009 (0.027)
Sanctity	-0.033 (0.047)	-0.005 (0.046)	-0.022 (0.044)
General	-0.010 (0.024)	-0.010 (0.023)	-0.004 (0.022)
Intercept	0.018 (0.024)	0.015 (0.023)	0.009 (0.022)
N	6304	6304	6304
Log-Likelihood	-4369	-4369	-4369

Table C.2: Logit models predicting argument persuasiveness as a function of moral congruence with OPs' opening statements (measured via cosine similarity in MFT dictionary results). Positive coefficients indicate higher probability of changing the OPs' mind ( $\Delta$  awarded). Standard errors (clustered by discussion thread) in parentheses. Estimates are used for Figure 4.7 in the main text.

Variable	Full Response Path	Root Response	Truncated Root Response
Moral Congruence	0.290 (0.056)	0.188 (0.056)	0.019 (0.054)
Intercept	-0.147 (0.028)	-0.092 (0.027)	-0.008 (0.024)
N	6304	6304	6304
Log-Likelihood	-4361	-4366	-4370