

# Morals Matter, But Not For Everyone\*

## The Conditionality of Moral Foundations in Political Reasoning

Patrick W. Kraft<sup>†</sup>

October 11, 2016

### Abstract

Moral Foundations Theory (MFT) proposes that political attitudes and ideology are structured by fundamental moral intuitions. While there is growing empirical support for the theory in the social sciences, recent studies have questioned the notion that moral intuitions can be viewed as stable traits. The present paper contributes to the literature by introducing an alternative measurement approach to explore the conditions under which moral foundations are expressed in politics. We utilize the Moral Foundations dictionary to identify references to moral intuitions when individuals discuss their political preferences in open-ended responses. The reliance on moral foundations in political reasoning is largely consistent with MFT, but whether and how individuals evoke moral considerations is conditional on their political knowledge and exposure to political discourse. Moral foundations have a powerful influence on public opinion and political behavior but this relationship is most evident among the politically knowledgeable—implying that moral reasoning is as much learned from the political world as it is innate.

**Keywords:** Moral Foundations Theory, Political Reasoning, Ideology, Open-ended Survey Responses

---

\*An earlier version of this paper was presented at the 73rd Annual Conference of the Midwest Political Science Association, April 16-19, 2015. The manuscript and code are available on GitHub: <https://github.com/pwkraft/mft>.

<sup>†</sup>Ph.D. Candidate, Stony Brook University, [patrick.kraft@stonybrook.edu](mailto:patrick.kraft@stonybrook.edu).

To what extent do the public’s political beliefs have a meaningful structure? This question has been of frequent scholarly interest in political science and related disciplines, yielding an array of different perspectives. Early accounts emphasized how ordinary citizens lack consistent political attitudes and knowledge necessary to form meaningful ideologies (Converse, 1964). However, with increasing levels of polarization and partisan sorting in contemporary politics (e.g. Mason, 2015), there has been renewed interest in systematic psychological and attitudinal differences between liberals and conservatives (Jost, 2006). One such area of research focuses on the differing moral concerns of liberals and conservatives. According to *Moral Foundations Theory* (MFT), moral thinking is organized by five innate intuitions: harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity (Haidt and Joseph, 2008; Graham et al., 2011). Liberals and conservatives differ in their relative emphasis of these foundations, with liberals prioritizing the foundations of harm/care and fairness/reciprocity, and conservatives endorsing all five dimensions more or less equally (Graham, Haidt, and Nosek, 2009).<sup>1</sup>

A series of studies builds on MFT and provides evidence for the influence of moral foundations on issue preferences (Koleva et al., 2012; Kertzer et al., 2014; Low and Wui, 2015; Clifford, Jerit, Rainey, and Motyl, 2015), candidate trait evaluations (Clifford, 2014), and vote choice (Iyer et al., 2010; Franks and Scherr, 2015). However, recent research cast some doubt on the theory’s basic tenets (e.g., Suhler and Churchland, 2011). Smith et al. (2016) empirically investigated three implicit assumptions underlying MFT, namely that moral foundations are individual traits that are relatively stable over time, that changes in moral foundations predict changes in political ideology, and that the foundations are heritable. The empirical evidence provided by Smith et al. (2016) does not support these assumptions, but this critique stands in contrast to the growing body of work that documents the political relevance of moral foundations. The present study attempts to reconcile this contradiction

---

<sup>1</sup>Subsequent accounts of MFT discussed the inclusion of further dimensions, such as *Liberty/Oppression* (c.f., Graham et al., 2013; Haidt, 2012). However, the analyses presented here will only focus on the dimensions initially suggested in Haidt and Joseph (2008).

by taking a different tack and directly examining the conditions under which the moral foundations are consequential to a person’s political judgments. Indeed, [Smith et al. \(2016\)](#) conclude that their findings are “consistent with the notion of moral foundations being highly responsive to context: more of a state than a trait.” Yet, few studies systematically investigated the conditionality of the relationship between moral foundations and political attitudes (see [Clifford, Jerit, Rainey, and Motyl, 2015](#), for a notable exception).

The present study provides a first step in this direction by directly analyzing the degree to which individuals vary in the expression of moral concerns in political reasoning and judgment. Using data from the 2012 American National Election Study, we create a measure of moral reasoning based on individual responses to the open-ended likes-dislikes questions and the moral dictionary introduced by [Graham, Haidt, and Nosek \(2009\)](#). Relying on open-ended questions rather than traditional survey-based measures of moral foundations allows us to investigate patterns of moral reasoning among individuals where the potential connection between morality and politics is not induced or facilitated by design.

The analyses reveal systematic differences between liberals and conservatives in their reliance on specific moral considerations when evaluating political parties and candidates even without being cued to think about morality. These differences, in turn, affect political evaluations as well as voting behavior—even after controlling for a person’s party identification. Most importantly, the overall reliance on moral foundations in political judgment as well as the respective differences between liberals and conservatives are conditional on knowledge and exposure to political discourse. Even though the ideological differences are broadly consistent with MFT, moral reasoning in the realm of politics is more context-specific than previous research suggests. From a methodological standpoint, the present study showcases techniques to improve the analysis of open-ended survey responses, such as automatic pre-processing, spell-checking, and coding of responses, while additionally correcting for potential deficiencies in the dictionary.

# Moral Foundations Theory

Moral values serve as a source for coherence in political attitudes and they shape individual belief systems. As such, one major focus of the moral foundations framework is the relationship between moral values and political belief systems (c.f., [Haidt, 2012](#)). Liberal morals have been shown to focus on “individualizing” foundations, which include harm/care and fairness/reciprocity. Conservatives, on the other hand, also emphasize the remaining foundations of ingroup/loyalty, authority/respect, and purity/sanctity, labeled the “binding” foundations ([Haidt and Graham, 2007](#); [Graham, Haidt, and Nosek, 2009](#)). Haidt and colleagues developed the Moral Foundations Questionnaire (MFQ) to measure individual differences in the emphasis of each foundation. The MFQ consists of a series of items that explicitly ask people to rate the relevance of different considerations when making decisions about right and wrong. The questionnaire also asks individuals to indicate their level of agreement with statements that represent the values implied by the five foundations.

Relying on the MFQ, subsequent research extended the initial findings. For example, scholars showed that moral concerns predicted attitudes towards a wide variety of divisive political issues (e.g. [Koleva et al., 2012](#); [Kertzer et al., 2014](#); [Low and Wui, 2015](#)). [Federico et al. \(2013\)](#) linked moral foundations to individual social dominance orientation (SDO) and right-wing authoritarianism (RWA). Further research directly investigated the relationship between moral foundations and candidate preferences ([Iyer et al., 2010](#)) or trait inferences about candidates ([Clifford, 2014](#)). Moral foundations have also been shown to predict turnout ([Johnson et al., 2014](#)) as well as voting behavior in the 2012 US Presidential election ([Franks and Scherr, 2015](#)). Overall, this research strongly supports the view that liberals and conservatives endorse different moral foundations and that these differences are related to political attitudes, evaluations, and behavior.

[Smith et al. \(2016\)](#), on the other hand, used the MFQ to examine some basic assumptions underlying MFT using data from several twin panels. The authors investigated whether moral foundations can be viewed as stable traits that vary little over time, whether changes

in moral foundations predict changes in ideology (or vice versa), and whether there are sufficient signs of heritability in the emphasis of moral foundations. Overall, the evidence reported in [Smith et al. \(2016\)](#) seems to contradict MFT: the foundations are less stable than ideology, do not seem to predict changes in the latter over time, and do not seem to be inherited through genes. However, it should also be noted that the analyses presented by [Smith et al. \(2016\)](#) are not without limitations. For example, their analyses of stability over time relied on different sub-scales of the MFQ in each of the waves, which might induce some of the instability between waves in the sample. Nevertheless, the results raise the possibility that the relationship between ideology and moral foundations might not be as stable as initial research suggested.

In order to reconcile the divide between the growing body of research highlighting the influence of the MFT on political attitudes and work challenging those conclusions, we need to directly investigate the conditions under which moral reasoning occurs when people think about politics. As we will argue below, this step involves measuring moral foundations in political reasoning without relying on the MFQ, which has been used in nearly all existing work (but see [Clifford, 2014](#)).

## The ‘Missing Link’ to Political Reasoning

Previous research has relied almost exclusively on MFQ, which explicitly asks people to judge the relevance of particular moral considerations (e.g. [Graham et al., 2011](#)). However, by directly asking people about the importance of considerations related to the five foundations, previous analyses presuppose an important link that requires more careful empirical investigation. More specifically, it does not allow us to directly examine the conditions under which the connection between moral values and political ideology manifests itself when citizens reason about politics and evaluate political actors. Measuring moral foundations using the MFQ therefore imposes constraints on our theoretical conceptualization of moral

reasoning and the range of hypotheses we can investigate. Some scholars have criticized the MFQ because it does not ask people to make moral judgments per se (e.g. Clifford, Iyengar, Cabeza, and Sinnott-Armstrong, 2015). Indeed, Graham, Haidt, and Nosek (2009, 1031) describe the reports on moral relevance as “self-theories of moral judgment,” rather than direct measures of judgment itself. Such abstract self-theories might, in turn, deviate from actual judgments in specific situations.

From a theoretical perspective, moral foundations are viewed as stable traits that affect attitudes and preferences regardless of potential individual differences and contextual effects. However, the incorporation of moral considerations in a specific political context could be much more variable. Previous research suggests that campaigns and elite communication can have important influences on individual moral reasoning. For example, Clifford and Jerit (2013) found that at the elite level, proponents and opponents of stem cell research place distinctive weights on moral foundations which in turn affected the public attitudes and the underlying considerations related to the issue. A subsequent study showed that elite rhetoric plays an important role in linking individual moral foundations with political attitudes, but only for those individuals who were most likely be exposed to it (Clifford, Jerit, Rainey, and Motyl 2015; also see Day et al. 2014). While these studies do not contradict MFT, they cast some doubt on the notion that moral reasoning in politics is as a direct reflection of stable moral intuitions.

The literature on *moral convictions* further suggests that moral reasoning in the realm of politics can be more variable than proposed by MFT. Skitka and colleagues argue that individuals hold moralized attitudes if they have a universal perception of “right and wrong” connected to the issue at hand (Skitka, Bauman, and Sargis, 2005; Mullen and Skitka, 2006; Skitka, 2010). This view implies that there can be considerable variance in the degree to which moral considerations are raised between individuals and across issues. However, the MFQ measures moral foundations as general traits that are implicitly assumed to apply in any context.

The present study addresses this gap by examining whether people utilize the foundations in their day-to-day political reasoning (i.e., without being prompted by the language of a questionnaire). Moving beyond the MFQ and measuring moral reasoning as the explicit expression of foundations allows us to directly examine the degree to which political attitudes are infused by moral convictions. As such, we investigate whether differences in moral reasoning between liberals and conservatives manifest themselves in a more unobtrusive context.

Insofar as moral intuitions play a role in political reasoning, citizens should rely on the moral foundations when reporting their attitudes towards political actors, even if they are not explicitly asked to do so. Thus, the first step of the analyses will be to replicate the findings connected to MFT and ideology using open-ended survey responses to general questions about political preferences. We expect that liberals will be more likely to spontaneously mention the moral foundations of harm/care and fairness/reciprocity than conservatives when evaluating political parties and candidates. Conversely, conservatives will be more likely to emphasize moral foundations of ingroup/loyalty, authority/respect, and purity/sanctity than liberals. Moral reasoning measured through open-ended responses is further expected to be politically consequential. The differences in the degree to which each foundation is emphasized should directly affect political preferences and vote choice consistent with the propositions of MFT.

After replicating the basic findings in the MFT literature in a more unobtrusive survey context that does not evoke issues of morality, the present study contributes to the literature by clarifying the conditionality of the foundations in day-to-day reasoning. Individuals who are exposed to the political process and resulting elite communications are expected to be more likely to incorporate moral considerations in their political reasoning because they adopt the respective arguments from elite discourse. Individuals who are not engaged in politics, on the other hand, might focus on other considerations when forming their preferences. The extent to which individuals rely on moral considerations when evaluating political ac-

tors is therefore not necessarily stable among individuals and across contexts. Rather, the tendency to emphasize moral foundations may be contingent upon individual levels of political knowledge, media exposure and political discussions.<sup>2</sup> We expect that individuals who are more exposed to political discourse and more engaged in politics will be more likely to emphasize moral foundations when evaluating political parties and candidates than those who have less experience and are less politically engaged. Furthermore, these differences are expected to exacerbate the ideological divide in moral reasoning.

The degree to which political attitudes are moralized as well as the ideological differences in moral reasoning cannot be investigated by relying solely on the MFQ. Instead, it is necessary to differentiate between a person's tendency to endorse a moral *foundation* (as a stable trait), and moral *reasoning* as the actual reliance on specific moral considerations and arguments in specific political judgments. Previous research has largely neglected this difference.

## Data, Variables, and Model Specification

The analyses presented here are based on the 2012 American National Election Study, which contains two representative cross-sectional samples. One sample was conducted by computer assisted face-to-face interviews while the other sample is based on an internet panel group. Both samples are pooled in the analyses. While each consisted of a pre-election and a post-election wave, most items described below are drawn from the pre-election wave.<sup>3</sup>

The major dependent variables are based on open-ended questions in which respondents were asked to report what they *liked* and *disliked* about either presidential candidate as well as the Republican and Democratic parties. More specifically, respondents were asked to list anything in particular that they like/dislike about the Democratic/Republican party as well

---

<sup>2</sup>Other scholars contend that the reliance on basic values is not contingent on individual characteristics such as political sophistication (e.g. [Feldman and Zaller, 1992](#); [Goren, 2001, 2004](#); [Marietta and Barker, 2007](#)).

<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.



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. The responses to the eight open-ended like/dislike questions (evaluating both parties and both candidates) were aggregated for each individual and pre-processed by correcting spelling errors using an implementation of the Aspell spell checking algorithm in R ([www.aspell.net](http://www.aspell.net)).

Respondents were not included in the analysis if they failed to provide an answer to all open-ended items, or if the interview language was Spanish. Table B.1 in the appendix provides an overview over the number of omitted cases. About 4% of the interviews were held in Spanish and about 7% of the respondents did not provide any open-ended response. Furthermore, Figure B.1 in the appendix 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.

Our goal is to use the open-ended responses to identify the extent to which respondents relied on specific moral foundations when talking about their political attitudes and preferences. We use the dictionary created by [Graham, Haidt, and Nosek \(2009\)](#) that identifies references to specific moral foundations. Other studies in the domain of MFT relied on a similar dictionary to investigate moral considerations in news media coverage about stem cell research ([Clifford and Jerit, 2013](#)).<sup>4</sup> In the present research, we use the same dictionary to analyze open-ended survey responses. Words like “protect” and “suffer” indicate references to the harm/care foundation, “equality” and “tolerant” signal reasoning based on the fairness/reciprocity foundation, “patriot” and “betrayal” indicate reference to the ingroup/loyalty foundation, “honor” and “respect” signal considerations related to authority/respect, whereas “integrity” and “duty” indicate reference to the purity/sanctity foundation.

[Graham, Haidt, and Nosek \(2009\)](#) developed the word lists as a collection of general associations related to the foundations that could be applied in any context. However, some

---

<sup>4</sup>The dictionary provided by [Graham, Haidt, and Nosek \(2009\)](#) is also presented in Appendix B.

words might be problematic when applied to the open-ended responses considered here. In particular, certain words might be too ubiquitous in the realm of politics and candidate evaluations to be regarded as an unambiguous indicator for specific moral considerations. For example, the moral foundations dictionary includes “leader,” “duty,” and “honor” as signal words for the authority/respect dimension. Should all three words be considered as equally indicative of the respective foundation? After all, if respondents are asked about their attitudes towards both presidential candidates, they might be inclined to describe them as good or bad *leaders* irrespective of explicit moral considerations related to authority. If that is indeed the case, the term “leader” should appear more frequently in open-ended responses across all individuals.

Conventional dictionary-based methods consist of dichotomous indicators or simple counts of signal word occurrences for each dictionary category in each text or response. As such, the only way to address this problem using conventional methods would be to revise the dictionary and eliminate words that are deemed problematic. However, this leaves a lot of discretion to the researcher. Instead, we rely on techniques developed in the field of information retrieval to match search queries to relevant text documents (see Manning et al., 2008, for an introduction). Broadly speaking, when a search engine identifies the set of documents most relevant to a query, it increases the importance of keywords that appear in fewer documents since those terms carry more discriminative information.

We rely on the same logic to create a measure (referred to below as “similarity score”) that represents the degree to which a person’s open-ended response resembles Graham et al.’s (2009) dictionaries for each moral dimension, while increasing the importance of more informative or discriminatory words in the dictionary. Formally, “similarity” is the tf-idf cosine score between the dictionaries and a person’s response to the open-ended items on the ANES<sup>5</sup>. A technical discussion appears in Appendix C, but the measure has several

---

<sup>5</sup>The acronym tf-idf stands for “term frequency - inverse document frequency,” which refers to the rationale that the frequency of specific terms are weighted by the inverse of the frequency of occurrence across documents.

useful properties that we note here. The similarity score is a continuous measure that ranges between 0 (no overlap between dictionary and response) and 1 (response is identical to dictionary), and is independent of response length. The score increases if a response includes a larger proportion of words that appear in the dictionary. Most importantly, however, words that appear in nearly all open-ended remarks affect similarity less than the words mentioned only by few respondents (because ubiquitous words convey less information about differences across individuals).

Overall, the similarity score provides a correction for potential distortions due to sub-optimal terms in the dictionary while leaving its exact content outside of the researcher’s discretion. While the resulting measure theoretically ranges from 0 to 1, we only observe very low values since open-ended responses are naturally quite distinct from the raw dictionaries. Since the exact values do not have a straightforward interpretation, the measure for each moral dimension is rescaled to unit variance<sup>6</sup>.

Due to the fact that the similarity scores are bound at zero (i.e. none of the words in the dictionary appear in the response), individual response patterns are modeled via Tobit regressions for each of the moral foundations under consideration. The Tobit framework allows us to decompose the estimates into the effect on the probability of mentioning a specific foundation *at all* as well as the degree of emphasis on the foundation given that it was mentioned by a respondent (see McDonald and Moffitt, 1980, for details on the decomposition of Tobit model estimates).

The key independent variable used to predict the emphasis on each of the moral foundations, is *political ideology*. Respondents were asked to place themselves on a seven-point scale ranging from extremely liberal to extremely conservative. Since there is no reason to suggest that moderates should fall in between liberals and conservatives in terms of their moral foundations (i.e. that the relationship between ‘continuous’ ideological self-placement

---

<sup>6</sup>Technically, values on this variable describe the cosine of the angle between the tf-idf vector space representation of the dictionary and each open-ended response. Please refer to the Appendix C for a more detailed discussion.

and the likelihood to mention specific moral foundations is inherently linear), we constructed dichotomous variables indicating whether respondent identified as liberals, conservatives, or moderates.

Additional control variables included in the analyses are *church attendance*, *education* (college degree), *age*, *sex*, *race* (African American), survey mode (online vs. offline), 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 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 relevance and consequences of moral reasoning measured through open-ended responses, the similarity scores for each of the moral foundations are used as independent variables to predict political outcomes. The dependent variables considered here are *candidate* and *party evaluations* (measured as the respective feeling thermometer differentials), as well as *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). In addition to the controls discussed previously, these analyses include measures of *party identification*, which were recoded similarly to ideology.

The next set of analyses investigates whether the expression of moral considerations in political judgment is conditional on knowledge and exposure to political discourse. The factors that are therefore expected to be related to references to moral foundations include *political knowledge*, which was measured as the sum of correct answers to factual knowledge questions. The analyses also investigate the effect of *political media exposure* and the frequency of *political discussions* with friends and family members. Here, we not only look

at the influence on individual foundations, but also examine whether these factors influence *general* moral reasoning. This latter variable is measured as the sum of individual similarity scores across all MFT dimensions (rescaled to unit variance after summation), which can be interpreted as a aggregate measure of how much respondents emphasize any moral consideration in their responses.

Whenever appropriate, independent variables were rescaled to range from 0 to 1. Figure B.2 in the appendix provides histograms of all independent variables included in different stages of the analyses.

## Ideological Differences in Moral Reasoning

Figure 1 provides a first descriptive overview of moral reasoning in open-ended responses. The figure displays the proportion of respondents who mentioned words that were included in the five different moral foundations dictionaries as well as their 95% confidence intervals.<sup>7</sup> Since responses for each individual represent their likes and dislikes across all eight open-ended items, each proportion indicates the percentage 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.

The moral foundation most frequently mentioned is harm/care: About 40% of the respondents mentioned at least one word included in respective dictionary. The second most frequently mentioned moral foundation is authority/respect with about 35%. The proportion of respondents emphasizing ingroup/respect of fairness reciprocity is slightly lower with about 28% and 22%, respectively. Purity/sanctity, on the other hand, was almost never mentioned by any of the respondents. This finding is surprising, since other studies found the foundation to be an important predictor of divisive political attitudes (Koleva et al., 2012). This result suggests that the terms contained in the purity/sanctity dictionary might be too

---

<sup>7</sup>Note that the proportions are based on the subset of the sample that provided a response to at least one of the open-ended items, and for which the interview was held in English.

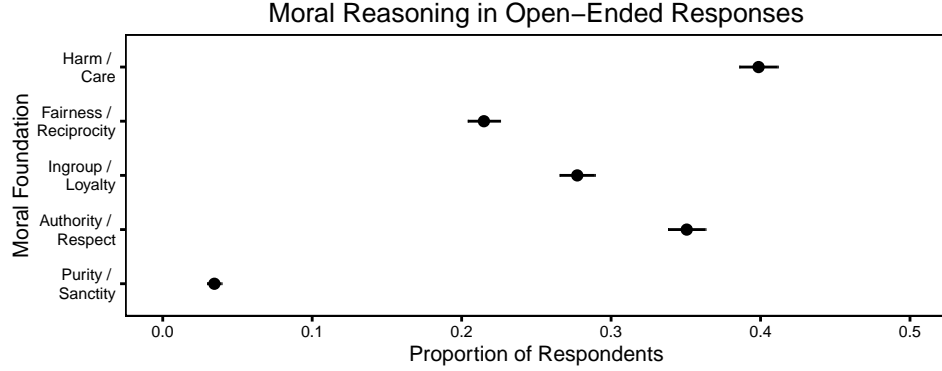


Figure 1: Proportion of respondents mentioning each of the moral foundations in any of their open-ended responses, along with 95% confidence intervals. The first two foundations are often labeled individualizing foundations, which have been shown to be more prevalent among liberals, while the remaining ones are described as binding foundations, which are more prevalent among conservatives.

uncommon in the context of politics and therefore not relevant for attitude expression. Due to the very rare mentioning of the purity/sanctity dimension, the subsequent analyses will concentrate on the remaining four moral foundations.<sup>8</sup> Subsequent analyses focusing on the dimension of purity/sanctity in open-ended survey responses might necessitate a revision of the moral foundation dictionary.

Overall, Figure 1 shows that a substantial proportion of individuals evokes moral considerations when describing their political attitudes even when they are not explicitly asked about morality. However, we are ultimately interested in ideological differences in the emphasis of moral foundations in open-ended responses. As such, we now turn to a more in-depth analysis of the similarity scores as measures of moral reasoning.

We begin by estimating a set of Tobit regressions using ideology (and the control variables discussed above) to predict the individual similarity score for each of the moral foundations (excluding purity/sanctity).<sup>9</sup> To reiterate, the similarity score measures the overlap between

<sup>8</sup>Unfortunately, this issue cannot not be properly addressed by relying on tf-idf cosine scores. The weighting can correct for some distortions due to individual ubiquitous terms in the dictionaries, but it cannot compensate for the fact that the purity dictionary as a whole contains mostly words that are never mentioned by respondents.

<sup>9</sup>The full estimates for this and all subsequent models discussed the remainder of the paper are presented in Appendix E.

individual open-ended responses and the dictionaries for each foundation and has a value of 0 if none of the words in a dictionary appear in a response. Figure 2 compares liberals and conservatives while holding all other variables constant at their respective means. The effects of ideology (liberal - conservative) estimated in the Tobit model are decomposed into two parts: the left panel displays the change in probability of the similarity score to be larger than zero (i.e. the probability of mentioning a specific foundation at all), whereas the right panel displays the expected change in the similarity score given that it is larger than zero (i.e. the degree of emphasis on a foundation given that it was mentioned, measured in standard deviations).

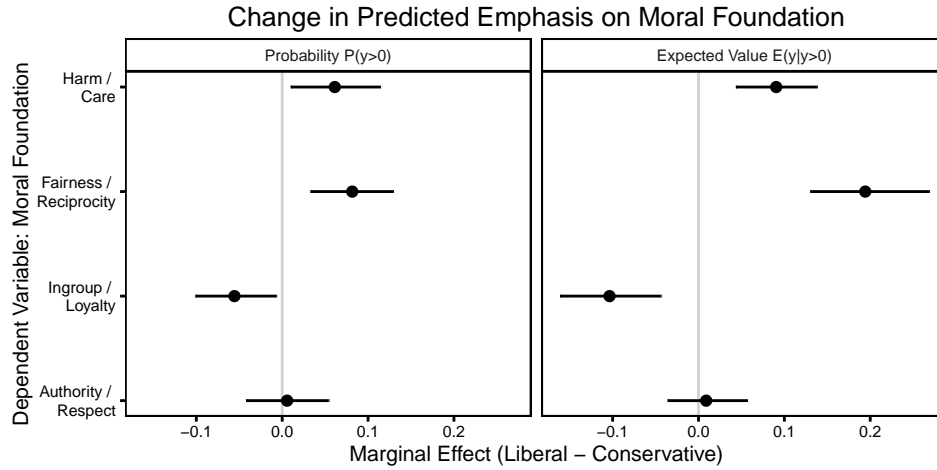


Figure 2: Difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel), and in the similarity score given that the foundation was mentioned (right panel), holding all other control variables at their respective means (along with 95% confidence intervals). Positive values indicate that liberals are more likely to mention the respective foundation than conservatives (left panel), or emphasize it more than conservatives (right panel), and vice versa. Estimates are based on separate Tobit models for each foundation's similarity score. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

Positive values denote a higher probability of mentioning the respective moral foundation (left panel) or a higher similarity score (right panel) of a response 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 significantly more likely to mention the foundations of harm/care and fairness/reciprocity. For respondents who identified as liberal as compared to conservative, the probability of referencing these two foundations is increased by about 10%. Furthermore, given that respondents mention these two foundations at all, liberals emphasize it more than conservatives when evaluating political parties and candidates. The similarity score for the harm/care foundation is about 0.1 standard deviations higher among liberals than conservatives. The effect is slightly larger for the fairness/reciprocity dimension. Conversely, being conservative increases the similarity score for the foundation of ingroup/loyalty by about 0.1 standard deviations. For the authority/respect dimension, we do not observe any significant differences between liberals and conservatives.

Taken together, the results are largely consistent with previous findings in the literature on MFT. Individuals evoke moral considerations when evaluating political parties and candidates without being explicitly asked about morality. Furthermore, there are systematic differences between liberals and conservatives in their reliance on binding and individualizing foundations (see [McAdams et al., 2008](#), for similar ideological differences when analyzing the content of life-narrative interviews). However, the fact that the authority/respect foundation showed insignificant patterns could suggest that the public’s reliance on moral foundations may be more context-specific than previously thought.

## The Political Relevance of Moral Reasoning

A skeptical reader might argue that even if the ideological patterns are consistent with MFT, the expression of moral considerations might not be as strongly related to other forms of political behavior as latent moral foundations measured by the MFQ. In order to alleviate this concern, we proceed to show that the expression of specific moral foundations in open-ended responses has important effects on political attitudes and behavior that match existing findings in the literature. Previous research relying on the MFQ has linked moral foundations



to an array of political outcomes, such as candidate preferences (Iyer et al., 2010) and voting behavior (Franks and Scherr, 2015). Do we see the same patterns for explicit moral reasoning as compared to latent moral foundations?

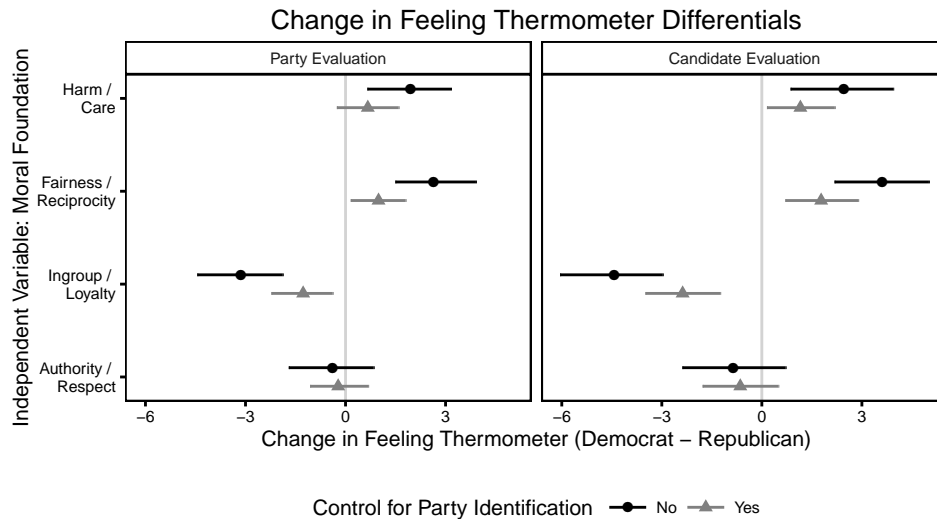


Figure 3: Change in predicted feeling thermometer differential when similarity score is increased from its minimum (no overlap between dictionary and response) by one standard deviation, holding all other control variables constant at their respective means (along with 95% confidence intervals). Positive 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 similarity scores for each foundation and gray triangles indicate estimates while additionally controlling for party identification. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

As a first step, we examine the relationship of moral reasoning and attitudes towards political parties and candidates. Figure 3 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 similarity scores for all moral foundations (including the control variables discussed above). 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 largely consistent with the previous results on ideological differences. Individuals who emphasize considerations

related to harm/care, and fairness/reciprocity 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 ingroup/loyalty dimension, they reported stronger preferences for the Republican party/candidate. These effects are robust after controlling for individual party identification. Thus, in both analyses in Figure 3, we observe sizable and significant effects for the influence of moral reasoning.

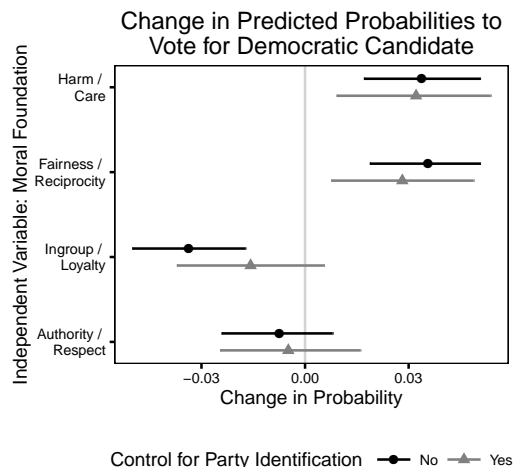


Figure 4: Change in predicted probabilities to vote for Democratic candidate when similarity score is increased from its minimum (no overlap between dictionary and response) by one standard deviation, holding all other control variables constant at their respective means (along with 95% confidence intervals). Positive values indicate that respondents who emphasize the respective foundation are more likely to vote for the Democratic candidate, and vice versa. Estimates are based on a single logit model including similarity scores or each foundation and dotted lines indicate estimates while additionally controlling for party identification. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

Figure 4 presents the changes in expected probabilities of voting for the Democratic (vs. the 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 similarity scores for each moral foundation as independent variables (as well as the remaining controls), which were held constant at their mean values when computing expected values. Again, the patterns are strikingly similar to the results presented

thus far, and largely consistent with MFT. Individuals who emphasized moral considerations related to the harm/care and fairness/reciprocity foundations are more likely to vote for Barack Obama than for Mitt Romney. Respondents who emphasized the ingroup/loyalty foundation, on the other hand, were less likely to vote for Obama.

The effects on political evaluations and vote choice might not seem large, but bear in mind that the measure of moral reasoning is solely based on open-ended responses, in which respondents were *not* explicitly asked about morality. The moral considerations evoked by respondents when describing their attitudes towards political actors were powerfully related to party and candidate evaluations as well as vote choice, even after controlling for strength of party identification and the length of individual responses. Overall, the analyses show that people’s open-ended comments about both candidates and both parties are imbued with moral content and that these comments relate to political judgments in the manner predicted by MFT.

## The Conditionality of Moral Reasoning

Having shown that liberals and conservatives differ with regard to the moral foundations they emphasize when evaluating political actors and that these differences allow us to predict preferences and vote choice, we now investigate whether the reliance on moral considerations varies between individuals as a product of political knowledge and exposure to political discourse. Campaign exposure and discussions influence the degree to which political attitudes are moralized to the extent that individuals adopt moral rhetoric from elites and their peers. As such, we first focus on determinants of the general tendency to emphasize *any* moral foundation by regressing the sum of similarity scores for each individual on political knowledge, political media exposure, and frequency of political discussions. Figure 5 depicts the respective effects when each independent variable is increased from its empirical minimum value to its empirical maximum value, holding all other variables constant at their means.

Again, estimates are based on tobit models that take into account the censoring of the moral reasoning measure and 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).

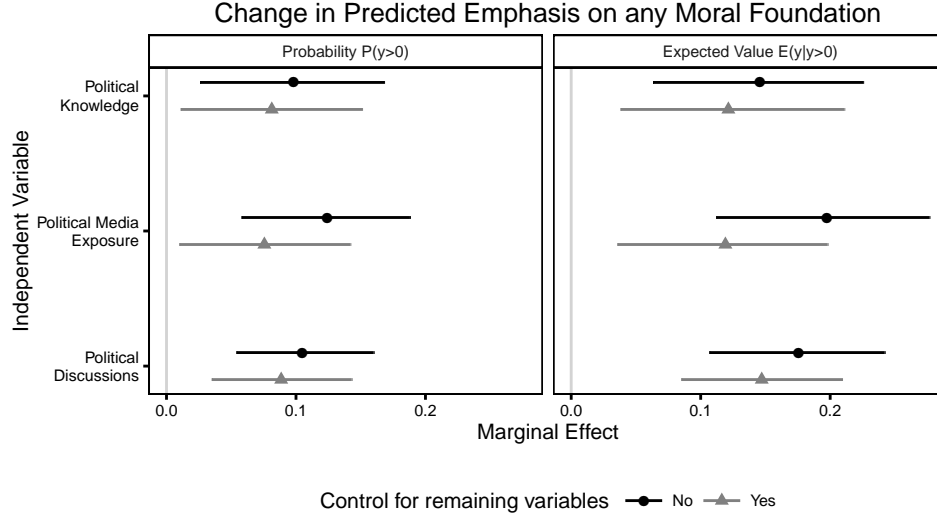


Figure 5: Change in predicted overall reliance on moral foundations depending on political knowledge, media exposure, and frequency of political discussions. The plot shows differences in predicted probabilities of mentioning any moral foundation (left panel) as well as in the summed similarity scores 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). Positive values indicate higher probability of mentioning, or stronger emphasis on moral foundations. Estimates are based on Tobit models and gray triangles indicate estimates while additionally controlling for the remaining variables presented in the figure. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

All variables considered here have a positive effect on the individual likelihood to mention as well as the respective emphasis on moral foundations when evaluating political parties and candidates. Higher political knowledge, higher exposure to political media and news, as well as more frequent political discussions increase the degree to which individuals rely on moral considerations. Thus, citizens *learn* to embed moral reasoning in their political evaluations. While moral intuitions themselves might well be innate, the extent to which

individuals make use of these intuitions when thinking about politics and evaluating political actors is context-dependent and subject to individual heterogeneity.

The significant positive effect of frequent political discussions (even after controlling for the remaining variables, political knowledge and media exposure), is especially interesting in this context. Citizens, who engage in frequent political arguments are more likely to use moral considerations when evaluating candidates and parties. This result suggests that morality serves as a rhetorical tool utilized to convince others of certain political views.<sup>10</sup>

Beyond their effects on *general* moral reasoning, we want to investigate whether the effects of political knowledge, media exposure, and discussions influence the systematic differences between liberals and conservatives in the emphasis on *specific* dimensions. Consider again the results presented in Figure 2, which showed that liberals were more likely to emphasize the harm/care and fairness/reciprocity dimensions, while conservatives were more likely to emphasize the ingroup/loyalty foundation. In order to investigate whether the differences are conditional on political knowledge, media exposure, and discussions, we extend the original models by including interaction terms for each of the three variables that have been shown to influence the overall reliance on moral considerations. Based on these models, we can examine whether the same ideological pattern can be recovered for individuals with low or high levels of political knowledge, media exposure, and discussion frequency.

Figure 6 displays the analyses equivalent to the previous results in Figure 2, now differentiating between individuals with high and low levels of political knowledge. While we can observe the same systematic differences between liberals and conservatives for individuals with high levels of political knowledge (upper panels), the effects of ideology disappear among individuals with low levels of political knowledge (lower panels) for all foundations except fairness/reciprocity. As such, we recover systematic differences in moral reasoning between

---

<sup>10</sup>This conclusion is also supported by the finding that the reliance on moral considerations is more pronounced among individuals who engage in non-conventional forms of participation (signing petitions, wearing campaign buttons, writing letters to representatives). On the other hand, simply participating in the previous election did not increase the emphasis of moral foundations significantly. These results are presented in the appendix, Figure D.1.

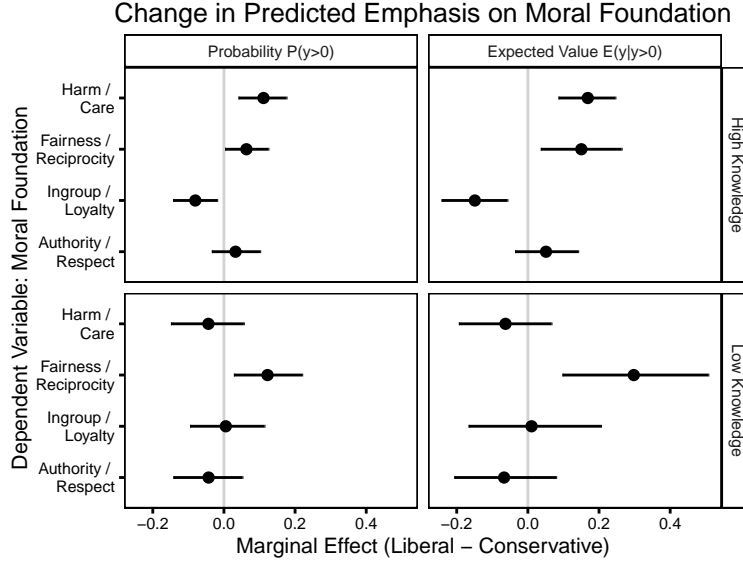


Figure 6: Difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel), and in the similarity score given that the foundation was mentioned (right panel), comparing individuals with low and high levels of political knowledge while holding all other control variables at their respective means (along with 95% confidence intervals). Positive values indicate that liberals are more likely to mention the respective foundation than conservatives (left panel), or emphasize it more than conservatives (right panel), and vice versa. Estimates are based on separate Tobit models for each foundation’s similarity score where ideology is interacted with political knowledge. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

liberals and conservatives only among the sub-group of highly knowledgeable respondents. It is worth noting again that all models presented here control for education, logged overall response length, and the Wordsum score, which should account for potential confounding factors related to the respondents’ eloquence when discussing their political attitudes. Of course, not all changes from significant to insignificant effects are themselves necessarily significant (c.f., Gelman and Stern, 2006). Figure D.2 in the appendix provides detailed results on the statistical significance of the difference-in-differences between the comparisons for this and the following analyses. In the case of political knowledge, we see that especially the ideological divide on the harm/care dimension is significantly larger among individuals with high political knowledge.

Next, we turn to a parallel analysis focusing on the influence of political media exposure

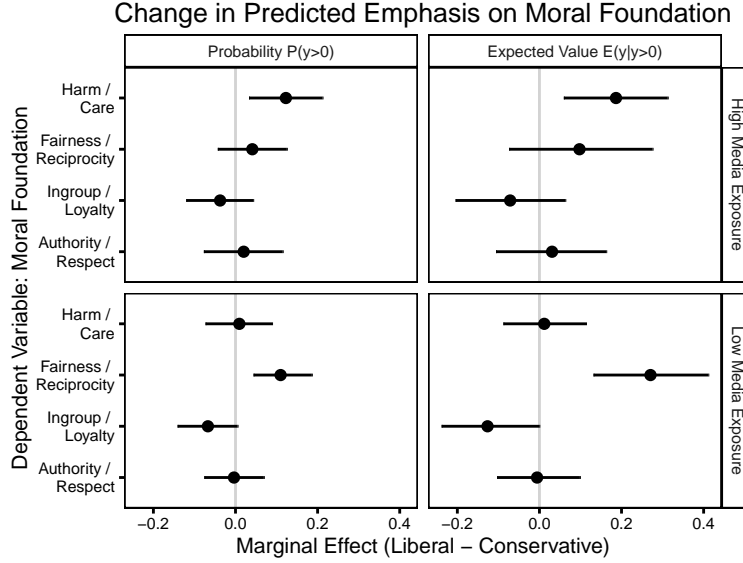


Figure 7: Difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel), and in the similarity score given that the foundation was mentioned (right panel), comparing individuals with low and high levels of political media exposure while holding all other control variables at their respective means (along with 95% confidence intervals). Positive values indicate that liberals are more likely to mention the respective foundation than conservatives (left panel), or emphasize it more than conservatives (right panel), and vice versa. Estimates are based on separate Tobit models for each foundation's similarity score where ideology is interacted with political media exposure. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

instead of political knowledge. Figure 7 compares the ideological differences for individuals with high and low levels of media exposure. Interestingly, the results look somewhat different than for political knowledge. While we again see that the higher emphasis on the harm/care foundation among liberals is only recovered among individuals who are frequent recipients of political news, the reverse is true for the fairness/reciprocity and ingroup/loyalty dimension. In both cases, the effects of ideology consistent with MFT only appear among individuals who are not frequently exposed to political news. However, only the change in the ideological divide on the harm foundation between high and low level of exposure reaches conventional levels of statistical significance (c.f., Figure D.2). Nevertheless, the result is interesting since it at least tentatively suggests that media consumption can decrease the ideological divide in moral considerations expressed in politics.

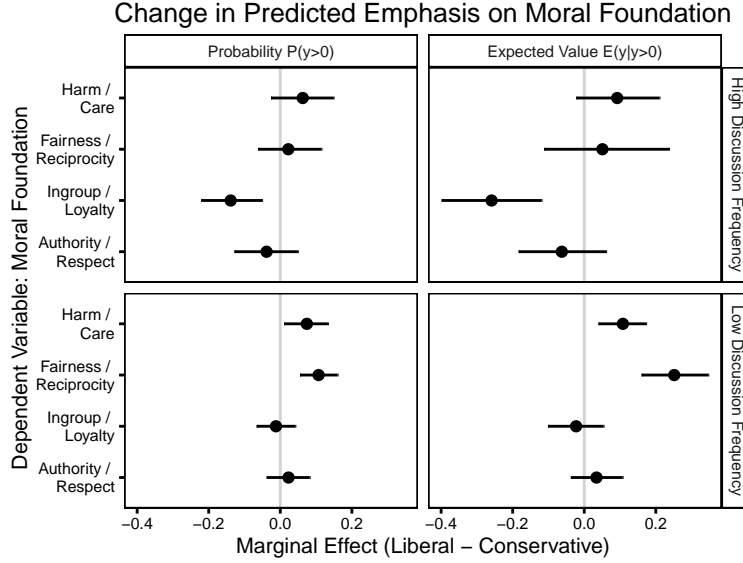


Figure 8: Difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel), and in the similarity score given that the foundation was mentioned (right panel), comparing individuals with low and high levels of political discussions while holding all other control variables at their respective means (along with 95% confidence intervals). Positive values indicate that liberals are more likely to mention the respective foundation than conservatives (left panel), or emphasize it more than conservatives (right panel), and vice versa. Estimates are based on separate Tobit models for each foundation’s similarity score where ideology is interacted with political discussions. The dimension of purity/sanctity was omitted due to its low general prevalence in individual attitude expressions (c.f., Figure 1).

Lastly, Figure 8 differentiates between individuals who discuss politics more or less frequently. Again, the results are not as straightforward as in the case of political knowledge. There are significant differences in the reliance on the harm/care and fairness/reciprocity dimension between liberals and conservatives who do *not* talk about politics with their peers, while no statistically significant effects are found among individuals who do discuss politics frequently. These changes between high and low discussion frequency, however, are not themselves statistically significant (c.f., Figure 9). On the other hand, the ideological divide in the emphasis on the ingroup/loyalty foundation is significantly increased among individuals who discuss politics frequently, such that it is only recovered among respondents who engage in frequent discussions.

These results draw much more nuanced picture of the ideological differences in moral



reasoning than previous studies. The patterns are only unequivocally consistent with MFT among politically knowledgeable respondents.<sup>11</sup> However, the picture is more complicated when looking at the influence of media exposure and political discussions. Exposure to political discourse increases the ideological divide on some dimensions but decreases it on others.

Explaining these patterns requires a more in-depth analyses of the content of the media coverage as well as the nature of political discussions. It might be too simplistic to assume that exposure to news media and discussions always exacerbate the ideological divide in moral reasoning. In fact, such a pattern should only occur if liberals and conservatives are exposed to rhetoric that raises the respective moral considerations. For example, if a conservative regularly watches news programs that emphasize moral arguments related to individualizing foundations, we would not expect him or her to be *less* likely to raise such considerations. In order to investigate this proposition, we make use of the fact that the 2012 ANES included items indicating whether individuals regularly watched a large array of news outlets. For all media sources available, we downloaded the complete content of the coverage during the last month of the campaign (October 2012) and coded their emphasis on moral foundations using the same approach as for open-ended survey responses. Figure B.3 in the Appendix provides an overview of the media outlets and their respective moral similarity scores for each dimension.

Based on the coded content for each media source, we now analyze whether individuals who mostly rely on media sources that emphasize a specific foundation more strongly were also more likely to mention the respective moral dimension in their open-ended responses. The results are presented in Figure 9. For each moral foundation, the independent variable consists of a dummy indicating whether the majority of media outlets used regularly by a respondent put a higher emphasis on the moral dimension than other sources (determined

---

<sup>11</sup>This is especially noteworthy given that many studies on MFT are based on opt-in samples collected through [www.yourmorals.org](http://www.yourmorals.org), which potentially leads to biases towards higher education among respondents.

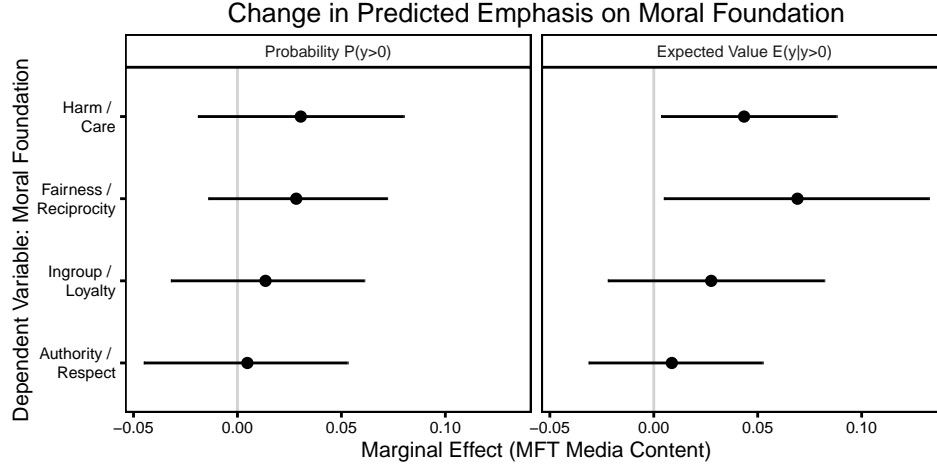


Figure 9: Content Analysis of Media Sources

by a median split in similarity scores). Note that the analyses therefore only incorporate the moral content of each media source rather than its respective ideological leaning or the respondent's ideology. The dependent variable is again the similarity score for each respective foundation. We can observe that individuals who are mostly exposed to media sources that discuss moral considerations of harm/care and fairness/reciprocity put a stronger emphasis on related moral arguments in their open-ended responses describing their political attitudes. Media exposure therefore induces moral reasoning consistent with the source's content.

As such, we would only expect media exposure to fosters the ideological divide in moral reasoning to the extent that the content of media sources exacerbates the respective patterns. This finding is quite encouraging given that [Haidt \(2012\)](#) prominently argued that incompatibility in moral foundations are an important source for disagreements in politics. The results presented here suggest that differences in moral reasoning can be overcome if individuals are exposed to similar rhetoric in political discourse. While there is no data available to examine the same argument for the case of political discussion, we view the preliminary results presented above as a strong case for subsequent investigations that incorporate the content of political discourse in the analysis of moral reasoning.

Overall, we can conclude from the analyses that whether researchers are able to recover systematic ideological differences in moral reasoning consistent with MFT is conditional on

individual-level factors. Political knowledge, media exposure and discussion frequency not only affect general levels of moral reasoning but also moderate the ideological gap between liberals and conservatives. Some portion of the ideological differences in emphasis on moral foundations can therefore be described as a product of learning in the political environment. However, the nature of the relationship between political context and individual moral reasoning is complex and requires further investigation.

## Examining Alternative Explanations

One of the greatest concerns for the analyses is whether the open-ended measure captures the theoretical concept of interest. It might be the fact that the ideological divide observed throughout the analyses does not necessarily reflect differences in expressed moral foundations. Rather, the 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 US Presidential election. For example, one important 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 order to address this concern, we replicated the main analysis using a different survey that varies the political context as well as the specific set of open-ended survey questions. It 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 Stony Brook University and included residents of both Suffolk and Nassau counties, New York. As such, the survey was conducted during a time without a major Presidential election campaign. 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 their own words. As such,

respondents are not asked to discuss specific political actors in a campaign, but rather describe their views on liberals and conservatives as social groups. The coding and analyses are equivalent to those for Figure 2, although the survey did not contain the Wordsum scores (and varying survey mode) included as controls in the main analyses. The results are displayed in Figure 10.

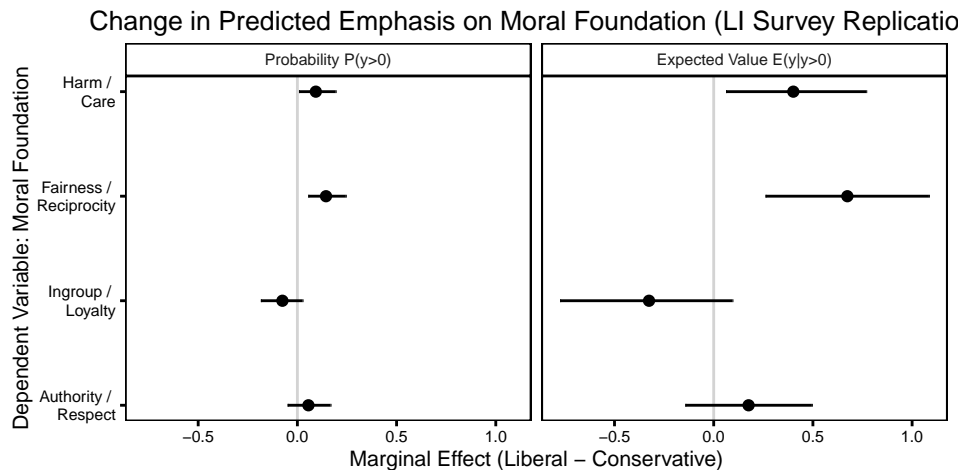


Figure 10: Replication of main model using Long Island Survey Data

The pattern is strikingly consistent with our previous results. While the result for the ingroup/loyalty dimension does not reach conventional levels of statistical significance, we find that liberals are more likely to emphasize the foundations of harm/care and fairness/reciprocity. Additional analyses reveal that the ideological differences are mostly due to the fact that individuals who identify as liberals emphasize the foundations of harm/care and fairness/reciprocity more strongly than conservatives when describing their ingroup (i.e., other liberals and their beliefs), while conservatives emphasize the foundation of ingroup/loyalty more strongly than liberals when describing their ingroup (results available upon request). The fact that we can recover the same basic ideological pattern in a survey that was conducted in a different political context (i.e., no Presidential campaign), and that relied on a different set of open-ended survey questions (asking about liberals and conservatives and their respective belief), provides further credence for the argument that the dictionary allows us to recover basic moral considerations in political reasoning .

## Conclusion

There is much more to be learned about the relationship between moral foundations and broader political attitudes and ideology (e.g., [Feldman, 2013](#)). To fully understand how political views are rooted in morality, researchers need to take a closer look at the conditions under which citizens express moral considerations in the context of politics. In contrast to previous accounts of MFT, we argue that the reliance on moral reasoning is moderated by political knowledge, media exposure, and political discussions. Not every individual necessarily thinks about politics in terms of morality. Rooting political preferences in deeper moral foundations requires exposure to a political discourse that fosters such a connection (c.f., [Clifford, Jerit, Rainey, and Motyl, 2015](#)).

This study utilized open-ended survey responses to investigate the conditionality of MFT in political reasoning. The analyses of open-ended survey responses is valuable in this context because it allows us to evaluate whether citizens make references to moral considerations in a political context that does not induce an explicit connection to morality. As such, we can directly investigate when and how ideological differences in the emphasis of moral foundations manifest themselves in individual reasoning about political actors. More generally, open-ended survey responses provide a promising and still largely neglected data source to investigate the determinants and structure of ideology and political reasoning. In particular, scholars can directly assess moral reasoning in surveys that do not contain the MFQ or related measures, simply by relying on open-ended survey responses.

The empirical analyses presented here extend and qualify previous research on moral foundations and ideology. We found systematic patterns in the emphasis on moral considerations among liberals and conservatives consistent with MFT for three out of four foundations. Liberals are more likely to mention considerations related to harm/care and fairness/reciprocity when discussing their political preferences, whereas conservatives are more likely to emphasize the moral foundation of ingroup/loyalty. The second part of the analyses focused on the political relevance of moral reasoning as conceptualized by open-ended survey responses.

Here, the results revealed consistent relationships between individual moral foundations and political preferences or voting behavior, which showed that moral reasoning (measured via open-ended survey responses) is a politically meaningful and influential concept. Lastly, political knowledge and exposure to political discourse increase the reliance on moral considerations. The evidence suggests that moral reasoning in political judgment is conditional on multiple individual-level factors and therefore likely to be part of a broader political learning process. At least in some cases, this learning process appears to imply an increased differentiation between liberals and conservatives in terms of the focus on specific foundations as described by MFT.

Moral foundations can provide a powerful source of structure in political belief systems. However, we cannot assume that the link between morality and politics is homogeneous across individuals and contexts. The present study therefore reaffirms the importance of moral reasoning in politics but highlights its conditionality. The results clarify instances when moral foundations are indeed influential in shaping ideological differences. Individual-level factors such as political knowledge, media exposure, and discussion frequency determine whether and how liberals and conservatives evoke moral considerations in political judgment. Ultimately, a deeper understanding of their influence necessitates further analyses of the broader political context that shapes individual information environments. As such, future research could examine how the endorsement of moral foundations varies over time and across campaigns. Such an investigation would allow us to further illuminate how exposure to political discourse fosters ideological differences in moral reasoning. A better understanding of the antecedents of this ideological divide is especially important in times of growing partisan polarization.

## References

- 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." *The Journal of Politics* 75 (03): 659–671.
- Clifford, Scott, Jennifer Jerit, Carlisle Rainey, and Matt Motyl. 2015. "Moral Concerns and Policy Attitudes: Investigating the Influence of Elite Rhetoric." *Political Communication* (forthcoming): 1–38.
- Clifford, Scott, Vijeth Iyengar, Roberto Cabeza, and Walter Sinnott-Armstrong. 2015. "Moral Foundations Vignettes: A Standardized Stimulus Database of Scenarios Based on Moral Foundations Theory." *Behavior Research Methods* forthcoming: 1–37.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: Free Press.
- 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.
- 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.
- Feldman, Stanley. 2013. "Political Ideology." In *The Oxford handbook of political psychology*, ed. Leonie Huddy, David O Sears, and Jack S Levy. Oxford University Press.
- Feldman, Stanley, and John Zaller. 1992. "The political culture of ambivalence: Ideological responses to the welfare state." *American Journal of Political Science* 36 (1): 268–307.
- 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.
- Gelman, Andrew, and Hal Stern. 2006. "The difference between significant and not significant is not itself statistically significant." *The American Statistician* 60 (4): 328–331.
- Goren, Paul. 2001. "Core principles and policy reasoning in mass publics: A test of two theories." *British Journal of Political Science* 31 (1): 159–177.
- Goren, Paul. 2004. "Political sophistication and policy reasoning: A reconsideration." *American Journal of Political Science* 48 (3): 462–478.
- 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.

- 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.
- Haidt, Jonathan. 2012. *The righteous mind: Why good people are divided by politics and religion*. Random House LLC.
- Haidt, Jonathan, and Craig Joseph. 2008. "The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules." In *The innate mind*, ed. P. Carruthers, S. Laurence, and S. Stic. Vol. 3. Oxford Univ Press.
- 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.
- 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.
- Johnson, Kate M, Ravi Iyer, Sean P Wojcik, Stephen Vaisey, Andrew Miles, Veronica Chu, and Jesse Graham. 2014. "Ideology-Specific Patterns of Moral Indifference Predict Intentions Not to Vote." *Analyses of Social Issues and Public Policy* 14 (1): 61-77.
- Jost, John T. 2006. "The end of the end of ideology." *American Psychologist* 61 (7): 651.
- Kertzer, Joshua D, Kathleen E Powers, Brian C Rathbun, and Ravi Iyer. 2014. "Moral support: How moral values shape foreign policy attitudes." *The Journal of Politics* 76 (03): 825-840.
- 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.
- Low, Michelle, and Ma Glenda Lopez Wui. 2015. "Moral Foundations and Attitudes Towards the Poor." *Current Psychology* forthcoming: 1-11.
- Manning, Christopher D, Prabhakar Raghavan, Hinrich Schütze et al. 2008. *Introduction to information retrieval*. Vol. 1 Cambridge university press Cambridge.
- Marietta, Morgan, and David C Barker. 2007. "Values as heuristics: core beliefs and voter sophistication in the 2000 republican nomination contest." *Journal of Elections, Public Opinion and Parties* 17 (1): 49-78.
- 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.



- McAdams, Dan P, Michelle Albaugh, Emily Farber, Jennifer Daniels, Regina L Logan, and Brad Olson. 2008. "Family metaphors and moral intuitions: how conservatives and liberals narrate their lives." *Journal of personality and social psychology* 95 (4): 978.
- McDonald, John F, and Robert A Moffitt. 1980. "The uses of Tobit analysis." *The review of economics and statistics*: 318–321.
- Mullen, Elizabeth, and Linda J Skitka. 2006. "Exploring the psychological underpinnings of the moral mandate effect: Motivated reasoning, group differentiation, or anger?" *Journal of personality and social psychology* 90 (4): 629.
- 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.
- Smith, Kevin B, John R Alford, John R Hibbing, Nicholas G Martin, and Peter K Hatemi. 2016. "Intuitive Ethics and Political Orientations: Testing Moral Foundations as a Theory of Political Ideology." *American Journal of Political Science*.
- Suhler, Christopher L, and Patricia Churchland. 2011. "Can innate, modular 'foundations' explain morality? Challenges for Haidt's moral foundations theory." *Journal of cognitive neuroscience* 23 (9): 2103–2116.

# Appendices

Online Appendices for:

Kraft, Patrick W. 2016.

“Moral Foundations of Political Reasoning. Investigating the Moral Underpinnings of Political Judgment.”

<b>A Moral Foundations Dictionary</b>	<b>35</b>
<b>B Additional Descriptive Information</b>	<b>36</b>
<b>C Description of Moral Reasoning Measure</b>	<b>39</b>
<b>D Additional Model Results and Robustness Checks</b>	<b>40</b>
<b>E Tables of Model Estimates</b>	<b>42</b>

# Appendix A Moral Foundations Dictionary

*Sources:*

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

*Note:*

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

## **Harm:**

safe\*, peace\*, compassion\*, empath\*, sympath\*, care, caring, protect\*, shield, shelter, amity, secur\*, benefit\*, defend\*, 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, unbiass\*, 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\*

## **Ingroup:**

together, nation\*, homeland\*, family, families, familial, group, loyal\*, patriot\*, communal, commune\*, communit\*, communis\*, comrad\*, cadre, collectiv\*, joint, unison, unite\*, fellow\*, guild, solidarity, devout\*, 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\*, immigrat\*

## **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\*, agitat\*, insubordinat\*, illegal\*, lawless\*, insurgent, mutinous, defy\*, dissident, unfaithful, alienate, defector, heretic\*, nonconformist, oppose, protest, refuse, denounce, remonstrate, riot\*, obstruct

## **Purity:**

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

## Appendix B Additional Descriptive Information

	N	Percent
Spanish Interview	228	3.86
No Responses	417	7.05

Table B.1: Missing open-ended responses

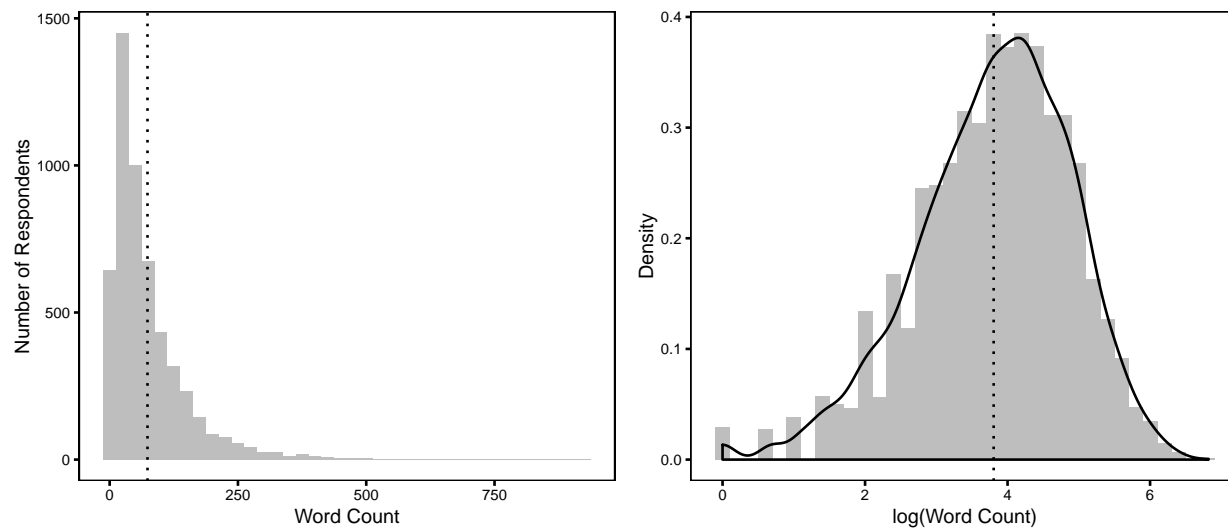


Figure B.1: Histograms displaying the distribution of individual response lengths in number of words for each respective item category. Dotted lines indicate the average response length.

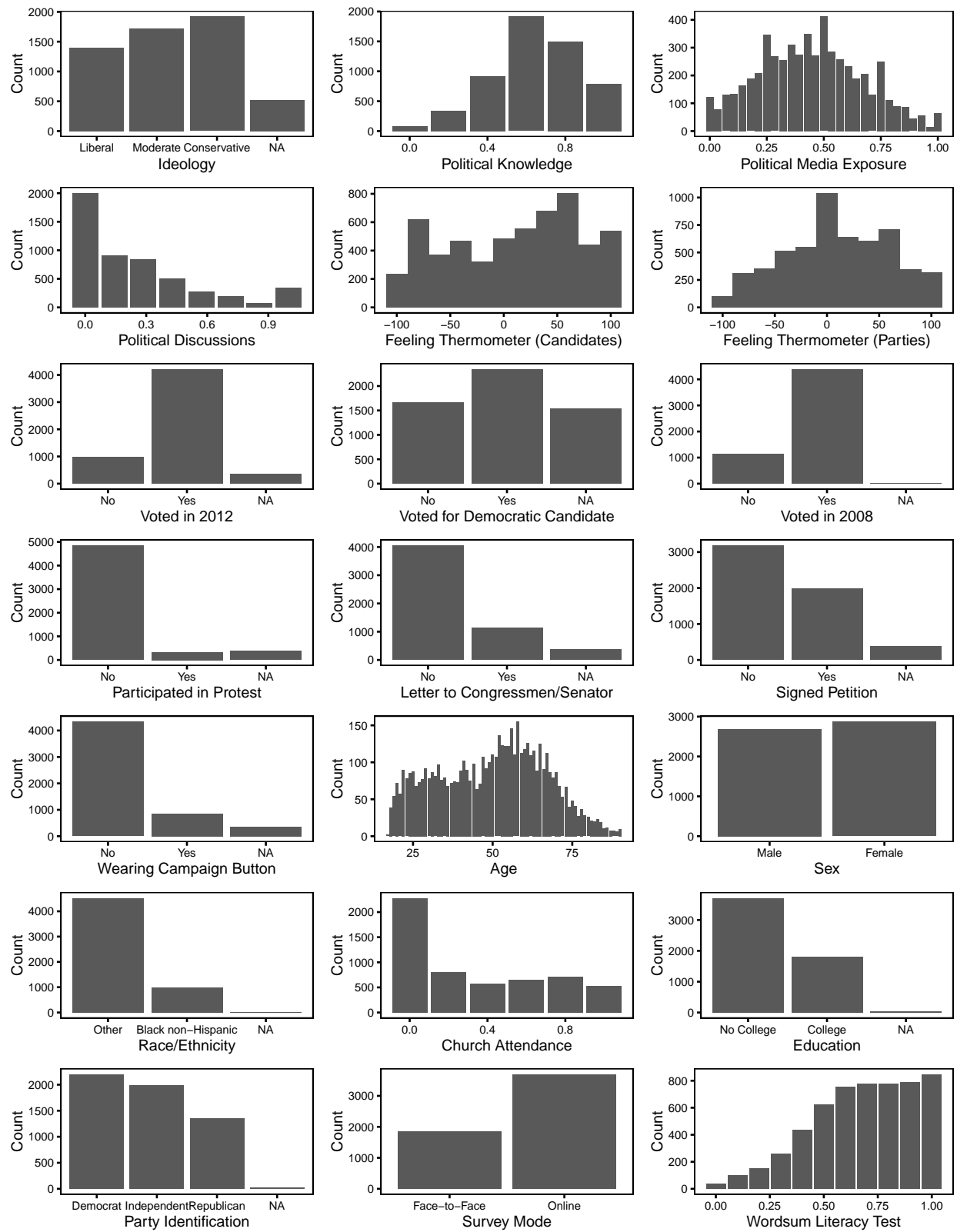


Figure B.2: Histograms for variables included in analyses.

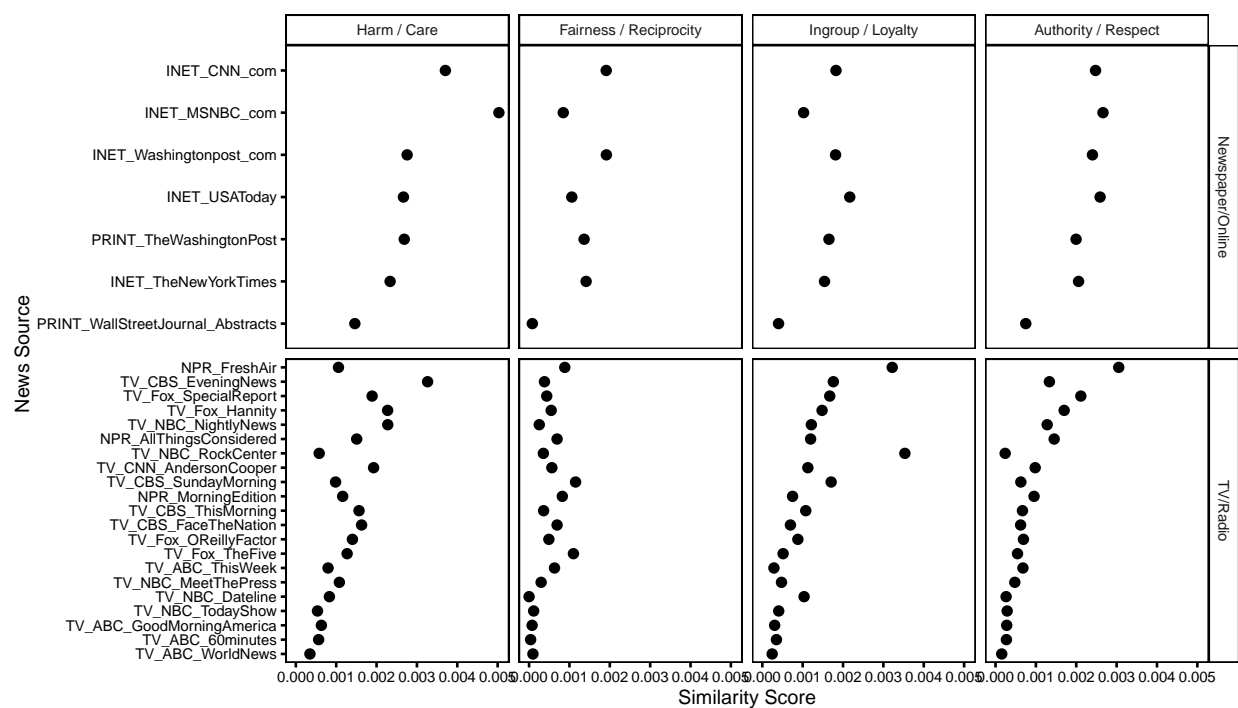


Figure B.3: Content Analysis of Media Sources

## Appendix C Description of Moral Reasoning Measure

This part of the appendix contains a detailed description of the measure of moral reasoning used throughout the paper. It is based on open-ended survey responses as well as the moral dictionary proposed by [Graham, Haidt, and Nosek \(2009\)](#). All techniques described here are adapted from common approaches in quantitative text analysis and information retrieval (see for example [Manning et al., 2008](#), for an introduction and from which much of the notation in this part of the appendix is adapted).

As a first step, each individual response as well as the moral foundations dictionaries are converted into a *document-term matrix*. Each row in the document-term matrix represents one document (i.e. the collection of open-ended responses for one individual, or the dictionary for a single moral foundation, respectively). Each column represents a term that is contained in the complete vocabulary of the corpus (i.e. all columns encompass the set of unique words contained in all documents). Each element in the matrix indicates the number of times the term is included in the respective document. As such, the matrix consists of raw frequencies of term  $t$  in document  $d$  (often denoted by  $tf_{t,d}$ ).

However, in the context of information retrieval that aims at matching string queries with most relevant matches out of a set of documents, term frequencies are usually weighted by the *inverse document frequency* in order to account for the fact that some terms have more discriminative power than others. The underlying logic is that a word that appears in almost all of the documents should be less decisive in affecting the similarity between the query and the document than a term that only appears in a selection of documents. The inverse document frequency is usually defined as

$$\text{idf}_t = \log \frac{N}{df_t}, \quad (1)$$

where  $N$  is the total number of documents in the corpus, and  $df_t$  denotes the number of documents in which term  $t$  occurs. The tf-idf score of term  $t$  in document  $d$  is then computed as

$$\text{tf-idf}_{t,d} = tf_{t,d} * \text{idf}_t. \quad (2)$$

Raw term frequencies are therefore weighted by the inverse of the term’s document frequency, which leads to increased (decreased) values if a term is only included in few (many) documents. We can now represent each moral dictionary and response as a vector of tf-idf scores for each term. For convenience, let  $w_{t,d}$  be the tf-idf score for term  $t$  in the dictionary and document  $d$  (which could be a moral dictionary or an open-ended response). Then,

$$\vec{m}_j = (w_{1,j}, w_{2,j}, \dots, w_{T,j}) \quad (3)$$

$$\vec{r}_n = (w_{1,n}, w_{2,n}, \dots, w_{T,n}), \quad (4)$$

where  $\vec{m}_j$  denotes the moral dictionary for foundation  $j \in \{1, \dots, J\}$ ,  $\vec{r}_n$  denotes the open-ended responses of individual  $n \in \{1, \dots, N\}$ , and  $T$  denotes the total number of unique terms in the entire text corpus. This representation of documents and queries (or in our case dictionaries) is usually described as the *vector space model* for representing the corpus.

Each moral dictionary and open-ended response is now represented as a vector of length  $T$ . A common measure of relevance, for example of a document ( $\vec{r}_n$ ) for a keyword search query (in our case  $\vec{m}_j$ ), is the cosine similarity:

$$\cos \theta_{j,n} = \frac{\vec{m}(j) \cdot \vec{r}(n)}{|\vec{m}(j)| |\vec{r}(n)|}, \quad (5)$$

where  $\vec{m}(j) \cdot \vec{r}(n)$  is the dot-product of the dictionary and the response, and  $|\vec{m}(j)| |\vec{r}(n)|$  is the product of the Euclidean norm of both vectors.  $\theta_{j,n}$  can therefore be understood as the angle between the vector space representation of the moral foundation dictionary  $j$  and the open-ended response of individual  $n$ . The cosine score is 1 if the relative frequency of words used by the respondent overlap completely with the terms in the dictionary (not that this holds irrespective of the length of the response), and 0 if there is no overlap.

## Appendix D Additional Model Results and Robustness Checks

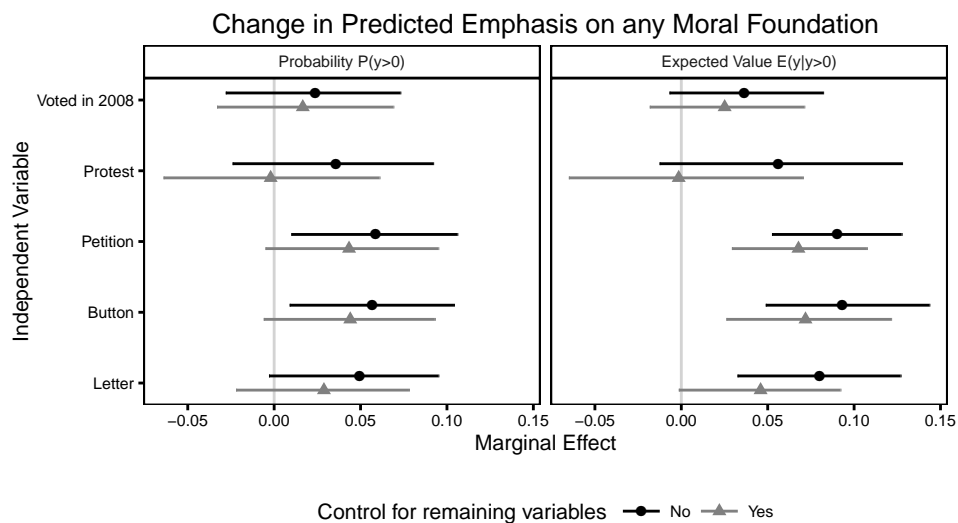


Figure D.1: Change in predicted overall reliance on moral foundations depending on previous turnout and non-conventional forms of participation (protest, petitions, campaign buttons, letter to congressmen/senator). The plot shows differences in predicted probabilities of mentioning any moral foundation (left panel) as well as in the summed similarity scores given that any foundation was mentioned (right panel), if a respondent engaged in the respective form of participation (vs. not) holding all other variables constant at their respective means (along with 95% confidence intervals). Positive values indicate higher probability of mentioning, or stronger emphasis on moral foundations. Estimates are based on Tobit models and gray triangles indicate estimates while additionally controlling for the remaining variables presented in the figure.



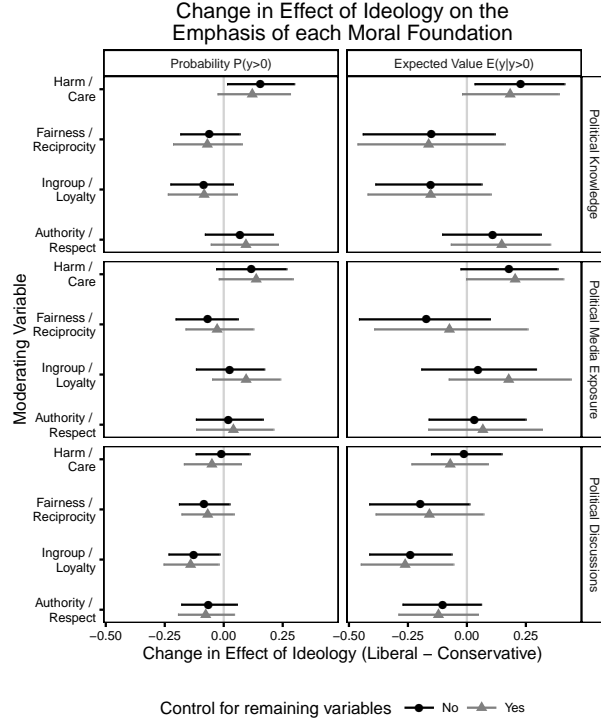


Figure D.2: Change in effect of ideology on emphasis of each moral foundation moderated by political knowledge, media exposure, and frequency of political discussions (difference-in-difference). The plot shows how the difference between liberals and conservatives in predicted probabilities to mention each moral foundation, as well as the respective similarity scores, change if each of the independent variables is increased from its minimum to its maximum value holding control variables constant at their respective means (along with 95% confidence intervals). Positive values indicate that liberals are more likely to mention a specific moral foundation if they score high on the moderating variable (knowledge, exposure, discussions, previous turnout, protest behavior), and vice versa. Estimates are based on individual Tobit models for each foundation and gray triangles indicate estimates while controlling for all remaining variables displayed in the figure.

## Appendix E Tables of Model Estimates