Measuring Morality in Political Attitude Expression[*]

Patrick W. Kraft[†]

March 5, 2017

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 conditional on the moral content of individual media environments.

Increasing levels of polarization has 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 (Haidt, 2012; Graham et al., 2013). Liberals and conservatives differ in their relative emphasis on these foundations, with liberals prioritizing the foundations of care and fairness, and conservatives endorsing all five dimensions equally (Graham et al., 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 whether moral foundations manifest themselves in a more unobtrusive context by examining how people utilize moral arguments in their day-to-day political reasoning without being prompted by the language of a questionnaire. Using a moral dictionary validated in previous studies (e.g., Graham et al., 2009), I propose novel methods 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. Insofar as moral intuitions play a role in political attitude expression, citizens should rely on the moral foundations when discussing their opinion 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, the analyses proceed to show that individuals who are exposed to moral rhetoric in political news are more likely to rely on moral considerations when discussing their political beliefs. More generally, I introduce methods to improve conventional dictionary-based approaches for the analysis of open-ended responses and showcase the integration of media content analyses to trace the influence of exposure to political discourse on individual response behavior.

Method

Using the moral foundations dictionary created by Graham et al. (2009), I identify references to specific moral considerations when respondents discuss what they like and dislike about political parties and candidates.[1] Other studies have relied on (variations of) this dictionary to investigate moral considerations in elite communication (e.g. Clifford et al., 2015), 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 et al., 2009). However, some terms in the dictionary might be problematic when applied to verbatim survey responses. In particular, when respondents describe their attitudes towards political actors, certain words might be too ubiquitous to be regarded as an unambiguous indicator for specific moral considerations. For example, the moral foundations dictionary includes "leader" as a signal word for the authority dimension. However, many respondents may be inclined to describe the qualities of presidential candidates as *leaders*

---

[1]See Appendix **??** for the full content of the dictionary.

irrespective of moral considerations related to authority.

One way to address this problem would be to revise the dictionary and eliminate words that are deemed problematic. Yet such revisions could be arbitrary and leave a lot of discretion to the researcher. Drawing on techniques developed in the field of information retrieval, I proposes an alternative approach. If it is the case that "leader" represents a term that is commonly used to describe presidential candidates (irrespective of moral considerations), it should appear more frequently in open-ended responses across individuals. 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. Stated differently, if a specific moral word is mentioned by a large majority of respondents, it is more likely that the term can be used in multiple contexts and is not necessarily unique to the moral domain. 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:

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

where 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 $\mathcal{D}_f$, and w_{it} denotes the number of occurrences of term t in document i. Furthermore, N denotes 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.[2] Terms that are ubiquitous across the entire corpus receive a lower weight, and terms that appear in only few documents receive a higher weight. The denominator in equation (1) includes +1 to ensure that it does not equal zero if a dictionary

---

[2]This specification is usually referred to as tf-idf weighting and is commonly used in quantitative text analyses. 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. See Manning et al. (2008, ch. 6) for an introduction.

term does not appear in any of the documents.

In the analyses presented here, 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 than the words mentioned only by few respondents because ubiquitous words convey less information about differences across individuals. Overall, the MFT score provides a correction for potential distortions due to suboptimal terms in the dictionary while leaving its exact content outside the researcher's discretion. Since nominal values of the MFT score above zero do not have a clear substantive interpretation, they are rescaled to unit variance.

Results

The analyses are based on the 2012 ANES. The primary dependent variables (i.e., the MFT scores described above) 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. The responses 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). The sanctity dimension was omitted from the analyses of open-ended responses due to its low general prevalence in individual attitude expressions. Due to the fact that the MFT 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. I 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 (c.f. McDonald and Moffitt, 1980). All models control for education, logged overall response length, and the Wordsum vocabulary score, which should account for potential confounding factors related to the respondents' eloquence when discussing their political attitudes.[3]

Ideological Differences

I begin by estimating a set of Tobit regressions using ideology to predict the individual MFT score for each of the moral foundations, which measures the weighted proportion of moral foundation terms in an open-ended response.[4] Figure 1 compares liberals and conservatives while holding all other variables constant at their respective means. The effects of ideology (liberal - conservative) are decomposed into two parts: the left panel displays the change in probability of mentioning a specific foundation at all (i.e., probability of the MFT score to be larger than zero), whereas the right panel displays the expected change in the degree of emphasis on a foundation given that it was mentioned (i.e., the change in the MFT score given that it is larger than zero, measured in standard deviations).
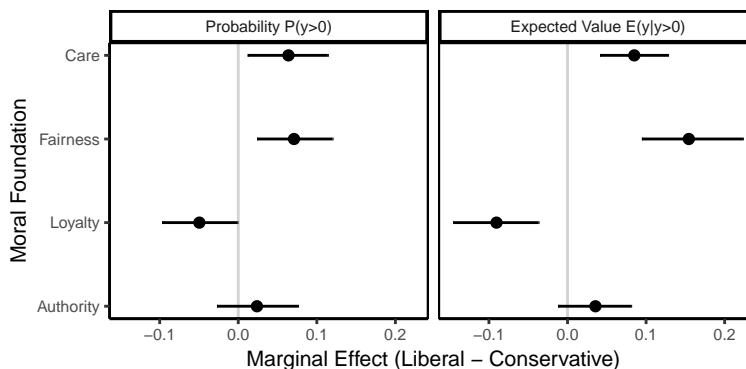

Figure 1: Difference between liberals and conservatives in the probability of mentioning each moral foundation (left panel) and in the MFT score given that the foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables include church attendance, education, age, sex, race, survey mode, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix.

---

[3]Please refer to the appendix for further information on the data and recoding.
[4]Full estimates for this and all subsequent models are presented in the appendix.

Positive values denote a higher probability of mentioning the respective moral foundation (left panel) or a higher MFT 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 the expectations of MFT for three out of four moral foundations. Liberals are significantly more likely to mention the foundations of care and fairness. More specifically, liberals were approximately 6 percentage points more likely than conservatives to reference these two foundations. 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 0.07 standard deviations higher among liberals than conservatives. The effect is slightly larger for the fairness dimension. Conversely, being conservative increases the MFT score for the foundation of loyalty by about 0.09 standard deviations. There are no significant differences between liberals and conservative on the authority dimension.

Moral Considerations and Vote Choice

A skeptical reader may 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 (e.g. vote choice) as latent moral foundations measured by the original MFQ. To address this concern, Figure 2 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 MFT 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. 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,

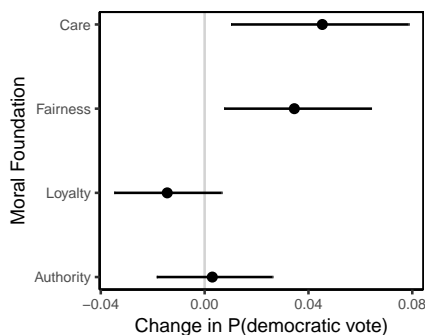on the other hand, were less likely to vote for Obama.[5]

Figure 2: Change in predicted probabilities to vote 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, church attendance, education, age, sex, race, survey mode, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix.

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. Yet, the moral considerations evoked by respondents are powerfully related to party and candidate evaluations as well as vote choice. 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.

Media Content and Exposure to Moral Rhetoric

Next, I investigate whether the general reliance on moral considerations is a product of exposure to moralized political discourse. 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 analyses of media sources used by each individual. Using Lexis-Nexis, I retrieved the content of 28 media sources covering

---

[5]A similar pattern can be observed in an analysis of feeling thermometers towards both parties and candidates. Please refer to the appendix for details.

either presidential candidate during the survey field period in the last month of the campaign (October 2012) and coded the general emphasis on any moral consideration using the same approach as for open-ended survey responses.[6] 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 aggregating the MFT scores of all media outlets watched or read by a given respondent.
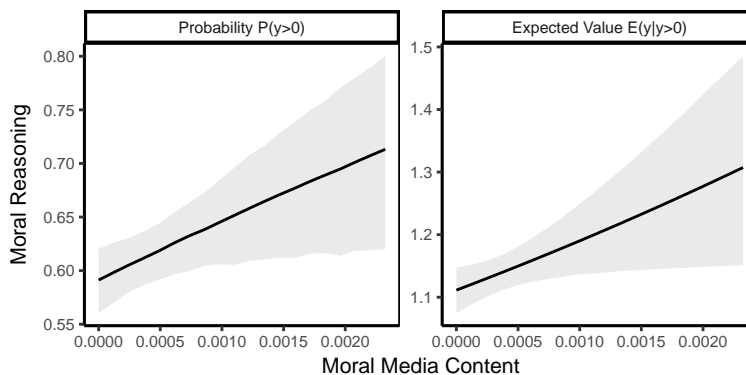

Figure 3: Effect of MFT content in individual media environments on the probability of mentioning any moral foundation (left panel), and on the summed MFT score given that any foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables include general media exposure, political knowledge, political discussion frequency, church attendance, education, age, sex, race, survey mode, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix.

Figure 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 arguments in their open-ended responses describing their political attitudes about the parties and candidates. Thus, citizens learn to embed moral reasoning in their political evaluations by adopting moral arguments from their media environment.

---

[6]Sources include for example the New York Times (print and online), CNN.com, or various Fox News Programs. See Figure **??** in the Appendix for more information.

Robustness Check

To this point, the analyses assume that the dictionary-based approach for open-ended responses captures the theoretical concept of interest—*moral* reasoning. Yet, the terms in the dictionary may be recovering other (i.e., non-moral) 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, a prominent issue in the 2012 election was the Affordable Care Act, which might increase the likelihood of Democrats mentioning the term "care" and thereby increasing the emphasis on the care foundation irrespective of underlying moral considerations. As such, observed differences between liberals and conservatives might be an artifact of the specific political context in which the survey took place.

To address this concern, I replicated the analysis from Figure 1 using survey data collected in a different political context. The survey 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 fielded during a major presidential election campaign. Furthermore, the survey uses a different set of open-ended items. Respondents were asked 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. The results are displayed in Figure 4.

The patterns 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 emphasize the loyalty foundation more strongly than liberals when describing their ingroup
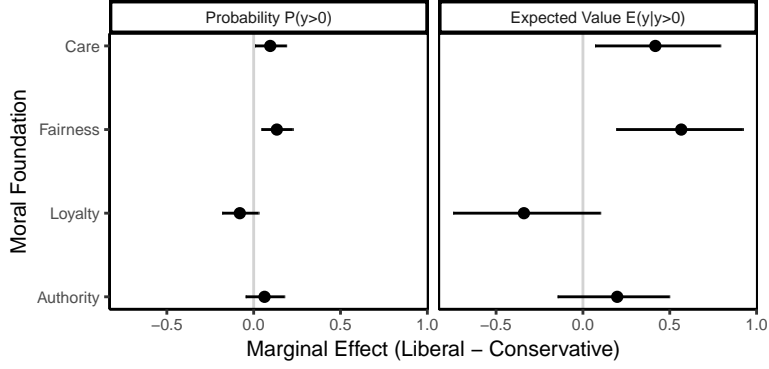
9

Figure 4: Replication of main model (c.f., Figure 1) using RDD sample from residents within a 25 mile radius of a large northeastern state university. 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.

(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.

Discussion

This study utilized open-ended survey responses to investigate morality in individual attitude expression. The analyses of open-ended survey responses is valuable because researchers can evaluate whether citizens make references to moral considerations in a context that does not induce an explicit connection to morality. Open-ended survey responses provide a promising and still largely neglected data source to investigate the role of morality in day-to-day political reasoning. Scholars can study moral reasoning in surveys that do not contain the MFQ simply by relying on open-ended survey responses. Furthermore, the approach outlined

here allows for a direct integration of the content of individual information environments in the analysis of political reasoning.

The empirical results presented here showed 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 care and fairness when discussing their political preferences, whereas conservatives are more likely to emphasize the moral foundation of loyalty. Moreover, there was a consistent relationship between individual moral foundations (measured via open-ended survey responses) and voting behavior. Lastly, exposure to moralized political discourse in the mass media increases a person's reliance on moral considerations.

References

Clifford, Scott. 2014. "Linking Issue Stances and Trait Inferences: A Theory of Moral Exemplification." *Journal of Politics* 76 (3): 698–710.

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.

Feinberg, Matthew, and Robb Willer. 2013. "The moral roots of environmental attitudes." *Psychological Science* 24 (1): 56–62.

Graham, Jesse, Brian Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter Ditto. 2011. "Mapping the Moral Domain." *Journal of Personality and Social Psychology* 101 (2): 366–385.

Graham, Jesse, Jonathan Haidt, and Brian 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 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.* New York: Random House.

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.

Jost, John. 2006. "The End of the End of Ideology." *American Psychologist* 61 (7): 651–670.

Kertzer, Joshua, Kathleen Powers, Brian Rathbun, and Ravi Iyer. 2014. "Moral Support: How Moral Values Shape Foreign Policy Attitudes." *Journal of Politics* 76 (3): 825–840.

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

McDonald, John F., and Robert A. Moffitt. 1980. "The Uses of Tobit Analysis." *The Review of Economics and Statistics* 62 (2): 318–321.