Measuring Morality in Political Attitude Expression*

Patrick W. Kraft†

March 7, 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.

*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.

†Ph.D. Candidate, Stony Brook University, patrick.kraft@stonybrook.edu.

Increasing levels of polarization have renewed scholarly interest in the psychological and attitudinal differences between liberals and conservatives (Jost, 2006). One such area of research focuses on the moral underpinnings of ideology. According to *Moral Foundations Theory* (MFT), moral thinking is organized by at least five dimensions: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation (Graham et al., 2013). Liberals and conservatives differ in their 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 how people utilize moral arguments in day-to-day political reasoning in a more unobtrusive context. Using a moral dictionary validated in previous studies (Graham 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 without being prompted by the language of a questionnaire. Insofar as moral intuitions play a role in political attitude expression, citizens should rely on the moral foundations when discussing their 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. Overall, this study introduces methods to improve conventional dictionary-based approaches for the analysis of open-ended responses and showcases the integration of media content analyses to trace the influence of exposure to political discourse on individual response behavior.

Method

This study utilizes the moral foundations dictionary created by Graham et al. (2009) to identify references to specific moral considerations when respondents discuss what they like and dislike about political parties and candidates.[1] Other studies have used (variations of) this dictionary to identify the moral foundations 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 dictionary terms are problematic when applied to verbatim survey responses. In particular, certain words might be too ubiquitous to be regarded as an unambiguous indicator for specific moral considerations. For example, "leader" is a signal word for the authority dimension. However, respondents may describe the qualities of presidential candidates as *leaders* irrespective of moral considerations related to authority.

One way to address this problem would be to revise the dictionary and eliminate problematic words. Yet such revisions could be arbitrary and leave too much discretion to the

[1] See the appendix for the full content of the dictionary.

researcher. Drawing on techniques developed in the field of information retrieval, I propose an alternative approach. If a specific dictionary term like "leader" is commonly used to describe presidential candidates, it is likely that the term can be used in multiple contexts and is not necessarily unique to the moral domain. Terms that are used by almost all respondents therefore provide less information about differences in their (moral) reasoning than terms that only occur in few responses. In this study, *MFT scores* are computed for a foundation by weighting each term in the dictionary according to its ubiquity across documents, which serves as a proxy for the term's discriminative information:

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

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 term does not appear in any of the documents.

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

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

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. Since nominal values of the MFT score above zero do not have a clear substantive interpretation, they are rescaled to unit variance.

Results

The open-ended responses in the 2012 ANES were aggregated for each individual and pre-processed by correcting spelling errors using the Aspell spell checking algorithm (www.aspell.net) before computing MFT scores. The appendix provides detailed information on the variables, recoding, as well as the raw proportions of individuals mentioning each foundation. Results for the sanctity dimension are not presented below due to its low general prevalence in individual attitude expressions.[3] To account for potential confounding factors related to the respondents' eloquence when discussing their political attitudes, all models reported below include controls for education, logged overall response length, as well as the Wordsum vocabulary score measuring verbal intelligence.

Ideological Differences

MFT scores measure the weighted proportion of moral foundation terms in an open-ended response. Since they are bounded at zero (i.e., response does not contain any moral words), I begin by estimating a set of Tobit regressions using ideology to predict individual MFT scores for each moral foundation.[4] Figure 1 compares liberals and conservatives while holding all other variables constant at their respective means. To facilitate the substantive interpretation, I decompose the estimates into the effect of ideology on the probability of mentioning a specific foundation *at all* (i.e., probability the MFT score is larger than zero) as well as the

[3] Only about 3.6% of respondents mentioned the sanctity dimension.
[4] Full estimates for this and all subsequent models are presented in the appendix.

degree of emphasis on the foundation given that it was mentioned by a respondent (i.e., the change in the MFT score given that it is larger than zero, measured in standard deviations).[5]

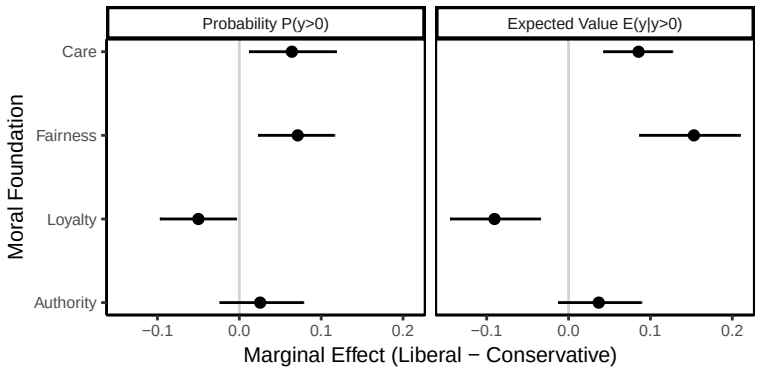

Figure 1: Difference between liberals and conservatives in the probability of mentioning a moral foundation (left panel) and in the MFT score given that the foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals). Control variables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

Positive values denote a higher probability of mentioning the respective moral foundation (left panel) or a higher MFT score (right panel) among individuals who identified as liberals, while negative values indicate a higher probability/higher score among conservatives. The effects are consistent with the expectations of MFT for three out of four moral foundations. Liberals are about 6 percentage points more likely than conservatives to mention the foundations of care and fairness. Furthermore, given that respondents mention these two foundations at all, liberals emphasize it more than conservatives when evaluating political parties and candidates. The MFT score for the care foundation is about 0.07 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.

[5]See for example McDonald and Moffitt (1980) for details on decomposing Tobit estimates.

Moral Considerations and Vote Choice

A skeptic may worry that the verbal expression of moral considerations might not be as strongly related to other forms of political behavior (e.g., vote choice) as moral foundations measured by the MFQ. To address this concern, Figure 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.[6]

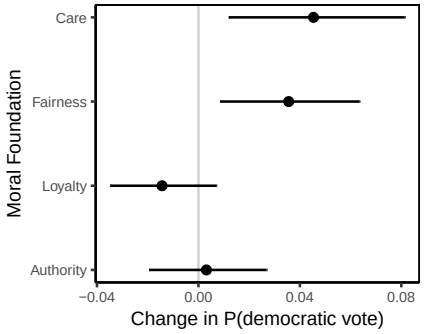

Figure 2: Change in predicted probabilities of voting for the Democratic rather than Republican candidate when MFT score is increased from its minimum (no overlap between dictionary and response) by one standard deviation, holding control variables constant at their respective means (along with 95% confidence intervals). Control variables include party identification, age, sex, race, church attendance, survey mode, education, response length, and the Wordsum vocabulary score. Full model results are displayed in the appendix.

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

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

respondents are powerfully related to party and candidate evaluations as well as vote choice. People's open-ended comments about both candidates and both parties are imbued with moral content that in turn relates to political judgments in the manner predicted by MFT.

Media Content and Exposure to Moral Rhetoric

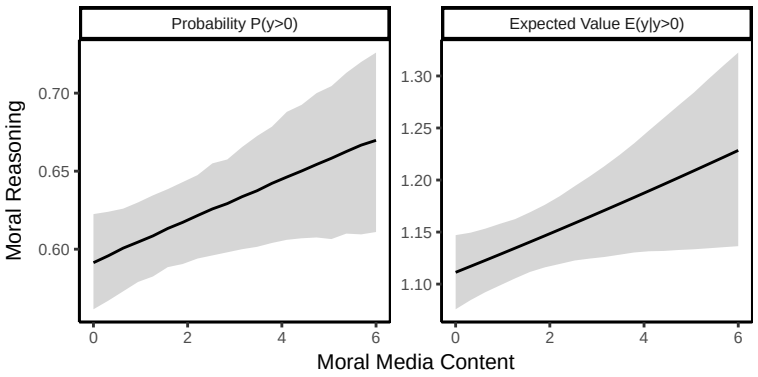

Figure 3: Effect of MFT content in individual media environments on the probability of mentioning any moral foundation (left panel), and on the summed MFT score given that any foundation was mentioned (right panel), holding control variables at their respective means (along with 95% confidence intervals).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

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 either presidential candidate during the survey field period in the last month of the campaign (October 2012) and coded the emphasis on moral considerations using the weighted dictionary approach described earlier. 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.[7]

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. Thus, citizens learn to embed moral reasoning in their political evaluations by adopting moral arguments from their media environment.

Robustness Checks

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. The appendix presents the results of multiple robustness-checks and additional analyses to alleviate this concern. For example, I replicate the analysis from Figure 1 using data collected in a different context (e.g., non-election year, Republican presidency, different set of open-ended items). The ideological differences in moral reasoning are consistent with previous results. A further robustness check compares the dictionary-based approach to measure moral media content with manual coding conducted by an independent group of researchers (Feinberg and Willer, 2013). Moralization measured using the dictionary is positively correlated with individual coder assessments of the degree to which news articles rely on moral rhetoric.

[7]Sources include for example the New York Times (print and online), CNN.com, or various Fox News Programs. Please refer to the appendix for more information on the media content analysis.

Discussion

Moral foundations theory has become an influential framework for understanding ideology and political attitudes. Yet, existing measures failed to directly assess whether individuals rely on moral considerations in their day-to-day political reasoning. This study filled this gap by examining moral arguments in individual attitude expression.

Consistent with MFT, the empirical results showed systematic patterns in the emphasis on moral considerations among liberals and conservatives for three out of four foundations. Liberals are more likely to mention considerations related to care and fairness when discussing their political preferences, whereas conservatives are more likely to emphasize the moral foundation of loyalty. Moreover, moral references in verbatim attitude expressions predicts voting behavior and the exposure to moralized political discourse in the mass media increases a person's reliance on moral considerations.

More generally, this study proposed improvements to conventional dictionary-based approaches in order to utilize a largely neglected data source: verbatim open-ended responses. Using this method, scholars can study moral reasoning in surveys that do not contain the MFQ simply by relying on open-ended items. Lastly, the approach outlined here allows for a seamless integration of media content in the analysis of moral reasoning.

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.

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.