ARTICLE

Lexical Ambiguity in Political Rhetoric: Why Morality Doesn't Fit in a Bag of Words

Patrick W. Kraft[1]* ⓘ and Robert Klemmensen[2]

[1]Juan March Institute and Department of Social Sciences, University Carlos III of Madrid, Spain and [2]Department of Political Science, Lund University, Lund, Sweden
*Corresponding author. E-mail: patrickwilli.kraft@uc3m.es

Abstract

How do politicians use moral appeals in their rhetoric? Previous research suggests that morality plays an important role in elite communication and that the endorsement of specific values varies systematically across the ideological spectrum. We argue that this view is incomplete since it only focuses on whether certain values are endorsed and not how they are contextualized by politicians. Using a novel sentence embedding approach, we show that although liberal and conservative politicians use the same moral terms, they attach diverging meanings to these values. Accordingly, the politics of morality is not about the promotion of specific moral values per se but, rather, a competition over their respective meaning. Our results highlight that simple dictionary-based methods to measure moral rhetoric may be insufficient since they fail to account for the semantic contexts in which words are used and, therefore, risk overlooking important features of political communication and party competition.

> *'Everyone cares about fairness, but there are two major kinds.*
> *On the left, fairness often implies equality, but on the right, it means proportionality.'*
> –Jonathan Haidt in *The Righteous Mind* (2012, 324)

David Easton famously defined politics as the 'authoritative allocation of values' (Easton 1953). It is no surprise that the resulting political debates ultimately centre around morality: questions of what is fundamentally right or wrong, This is especially the case in countries that are becoming increasingly polarized, such as the US or the UK. Consequently, recent scholarship in political science and beyond has developed a renewed interest in the relationship between politics and morality (Barker and Tinnick 2006; Lakoff 2010; Ryan 2014). A particularly influential framework on how to understand moral values related to political worldviews is the Moral Foundations Theory (MFT). The framework highlights that morality varies along at least five distinct dimensions: care, fairness, loyalty, authority, and sanctity (Haidt 2001, Haidt 2012; Haidt and Joseph 2004; Haidt, Graham, and Joseph 2009; Haidt, Koller, and Dias 1993). A consistent finding is that liberals and conservatives vary systematically in their endorsement of these five dimensions (see, for example, Clifford 2017; Clifford et al. 2015; Graham, Haidt, and Nosek 2009; Kraft 2018). While the causal direction of the relationship between morality and ideology is contested (Hatemi, Crabtree, and Smith 2019; Smith et al. 2017), the strong connection between political predispositions and moral values in the public suggests that these differences should

manifest in the context of elite political discourse. Specifically, politicians, party leaders, and candidates running for office should have incentives to make use of different moral arguments to target their supporters' underlying political convictions (Blumenau and Lauderdale 2022).

Despite the robust relationship between moral values and ideological differences in the general public, recent work examining the use of moral rhetoric in elite communication has yielded mixed results. While some scholars found differences between liberal and conservative politicians that were largely consistent with the expectations of MFT (for example, Brady et al. 2019; Clifford and Jerit 2013; Enke 2020), other studies found less conclusive evidence regarding ideological differences in moral rhetoric (for example, Frimer 2020; Sterling and Jost 2018; Wang and Inbar 2021). We argue that these empirical inconsistencies can be explained by two insights about the nature of moral messaging that have not been sufficiently accounted for in previous research: in brief, ideological differentiation in moral messaging is (a) conditional on the composition of the target audience that is addressed by elites and (b) can be more subtle than simply emphasizing one moral dimension over another.

We address the first gap in the literature by systematically comparing different types of political speeches that vary in the extent to which they address a wider national audience of mainly in-party supporters across two countries. In the US, we compare ideological differentiation in moral messaging between State of the Union Addresses as well as the party convention speeches of presidential candidates. We employ a similar contrast in the United Kingdom, where we analyze the moral rhetoric in the Queen's Speeches written by the British government (delivered by the Queen/King) in comparison to yearly party leader speeches. The State of the Union Addresses and the Queen's Speeches are intended for national audiences and are frequently evaluated by pundits in terms of their potential to unify the country, while speeches at party conventions primarily target the politician's supporters, often in an attempt to mobilize them for upcoming elections. Consequently, we expect that ideological differentiation in moral rhetoric will be more pronounced among the latter set of speeches in both countries.

However, this does not necessarily imply that there is no ideological differentiation in moral messaging when targeting national audiences, which leads us to the second gap in the literature we are going to address in this research. Standard dictionary approaches that only focus on whether certain values are endorsed are insufficient to detect more nuanced differences in moral rhetoric. Instead, we examine how moral arguments are contextualized by developing a sentence embedding approach that allows us to examine differences in the semantic contexts in which elites across the ideological spectrum employ moral language.

Our results indicate that while liberal and conservative politicians use similar moral terms, there is evidence that they attach systematically different meanings to the underlying values. Interestingly, this divergence is more pronounced for moral arguments focused on loyalty in the US and fairness in the UK. In addition, while moral divergence is evident across all types of speeches, it appears to be stronger when elites address their supporters rather than broader national audiences. We corroborate our findings in a set of two supplemental analyses. First, we examine rhetoric used in US presidential debates. This allows us to explore if ideological differentiation in moral reasoning persists in settings where elites target the same audience but have much less control over the issues that are debated and, importantly, both sides have to discuss the same topics. Second, we compare email newsletters from US senators. Here, we have an opportunity to assess divergence in moral rhetoric when politicians are directly communicating to dedicated supporters instead of a broader audience. In sum, we examine elites' moral rhetoric across a variety of contexts with a changing audience composition.

Our findings suggest that the politics of morality is not only about the endorsement of specific moral values per se but, rather, a competition over the meaning of such values. We illustrate the nuanced ways elites use moral language and thereby underscore the necessity to move beyond simple dictionary approaches, especially when it comes to the study of elite communication (see also Rheault and Cochrane 2020; Rodman 2020). Thus, in light of Haidt's (2012) basic insight that

moral foundations such as fairness can have diverging meanings, text-based approaches to studying elite communication need to incorporate the semantic context of moral rhetoric to help us better understand the politics of morality (see also Hochschild 1981; Meegan 2019).

Moral Foundations and Elite Rhetoric

Moral Foundations Theory has been successful in establishing systematic associations between moral foundations and political ideology across various contexts (Graham et al. 2013). Liberals put greater emphasis on values such as care and fairness while conservatives rely more on values such as loyalty, authority, and sanctity. The endorsement of these moral foundations has also been linked to more specific attitudes such as towards the poor (Low and Wui 2015), foreign policy (Kertzer et al. 2014), environmental protection (Feinberg and Willer 2013), contentious social issues like abortion (Koleva et al. 2012), and political preferences expressed in open-ended responses (Kraft 2018). While primarily tested in the US, these patterns generalize to other contexts, such as in Sweden (Nilsson and Erlandsson 2015), the Netherlands (Van Leeuwen and Park 2009) and South Korea (Kim, Kang, and Yun 2012). In sum, the political relevance of moral foundations at the individual level is well-established in the literature (see also Turner-Zwinkels et al. 2021).

The fact that citizens systematically endorse distinct moral values across the ideological spectrum creates incentives for political elites to leverage these differences as part of their communication strategies. Specifically, political messaging has been shown to be more effective if it relies on moral arguments consistent with people's predispositions. First, framing issues in terms of moral foundations strengthens prior attitudes (Day et al. 2014) and facilitates persuasion (Feinberg and Willer 2013, 2015). Second, moral appeals in political advertising induce emotional responses (Lipsitz 2018), and moral-emotional content is shared more widely on social media (Brady et al. 2019). Third, stronger moral convictions are associated with greater political engagement (Skitka and Bauman 2008). Accordingly, politicians should have strong incentives to rely on distinct moral rhetoric to mobilize their supporters, to persuade out-partisans, or to decrease support for their opponent (see also Blumenau and Lauderdale 2022; Voelkel and Feinberg 2018).

Several recent studies have, therefore, examined how politicians use moral language in their communication but evidence on the political elites is surprisingly mixed. On the one hand, some researchers report ideological differences in elite moral rhetoric that are consistent with MFT. Clifford and Jerit (2013), for instance, demonstrate how proponents and opponents of stem cell research rely on different moral arguments. Similarly, using latent semantic analysis, Sagi and Dehghani (2014) find ideological differences in moral arguments surrounding the debate over abortion in the US Senate, where Democrats rely more on fairness-related arguments and Republicans emphasize the purity foundation. Jung (2020), while not examining ideological differences on individual foundations, finds that moralized language in party manifestos across six English-speaking countries enhances voter turnout. Other studies, however, find less consistent evidence that politicians differ in moral language use, as suggested by MFT (for example, Frimer 2020; Wang and Inbar 2021). For example, Sterling and Jost (2018) find that liberal members of Congress used more fairness- and harm-related words in Twitter messages, while conservative legislators used more authority-related words, a finding that is consistent with MFT. At the same time, however, liberals used more language about group loyalty and purity, which is unexpected according to the theoretical framework. Moving beyond the US context, Bos and Minihold (2022) find similar evidence in European multiparty systems, suggesting that the political elites frequently use moral language, albeit not always in a way that is coherent with the predictions of MFT.

A possible explanation for these empirical inconsistencies can be found in the recent study by Enke (2020), who analyzed survey data of more than 200,000 American citizens, to show that there is systematic variation between the endorsement of moral foundations at the county level. Importantly, representatives in the U.S. Congress use rhetoric that aligns with their

constituency's moral preferences: comparing the campaign documents and congressional speeches of various candidates, the author finds that political elites are using moral language that is congruent with their electorate. Focusing on communications in the context of U.S. presidential elections, however, Enke (2020) also shows that ideological differences in moral messaging tend to dissipate after the conclusion of the primaries. Thus, there appears to be little evidence for moral divergence if political elites address the same audience.

In sum, these findings suggest that differentiation in moral rhetoric diminishes once politicians compete in front of the same audience. While these results suggest a clear link between elite-level moral rhetoric and specific audiences, the question remains: how do politicians address morally diverse audiences; for example, when elected elites run for national office? Is political rhetoric simply changing in nature and less infused with morality or do elites employ moral messaging that is still distinct but more nuanced? In the next section, we argue that politicians have incentives to send ambiguous signals when the moral preferences of their audience are unknown. Consequently, we need analytic tools that are capable of detecting more subtle cues in political rhetoric.

Moral Messaging and Audience Composition

Previous research has revealed mixed evidence regarding the systematic ideological differences in moral foundations among political elites. While some studies only suggest limited ideological differentiation, others report clear evidence that politicians adjust their messaging to match their respective audiences. We argue that these diverging results are due to the variation in intended audiences and the underlying differences in moral preferences. In a setting where variation in moral foundations is limited and preferences are known to the speaker, it is without risk for politicians to incorporate unambiguous moral signals. However, when there is greater variation in moral preferences or where the preference distribution is unknown, politicians are faced with a strategic dilemma: do they engage in clear moral rhetoric with the risk of being punished by voters or should they send equivocal moral signals?

Building on the idea that voters seek to minimize the moral distance between themselves and their representatives (Enke 2020), we argue that elites have an incentive to employ strategic ambiguity in moral messaging when they address broader audiences. Canonical work as early as Downs (1957) suggests that parties can have strategic incentives to send equivocal messages to enhance their electoral potential. Shepsle (1972) similarly argues that elites can have incentives to pursue a strategy of ambiguity, depending on certain characteristics of the electorate (see also Page 1976). In short, comparable to electoral incentives for equivocal messages on specific policy positions, different audience compositions may invoke equivalent incentives for ambiguity in moral messaging by political elites. Therefore, we formulate our first hypothesis as follows:

> *H1:* Ideological differentiation in moral messaging in political speeches is stronger if elites address their supporters than if they address broader national audiences.

As we point out above, however, moral divergence may be more subtle than simply emphasizing one moral dimension over another. Instead, moral messaging may incorporate competition over the meaning of shared values which, ultimately, invokes lexical ambiguity in political rhetoric. Thus, our second hypothesis directly addresses the extent to which political elites attach the same meaning to moral foundations or whether they use them in different semantic contexts:

> *H2:* Ideological differentiation in moral messaging in political speeches manifests itself through elites emphasizing the same moral foundations in different semantic contexts.

In sum, we expect to find more systematic ideological differentiation in elite rhetoric when politicians address their core constituents who themselves have distinct moral preferences. We further

argue that this does not necessarily imply that there is no moral dimension to political discourse when addressing broader national audiences. However, instead of focusing on whether certain values are endorsed, we have to examine how they are contextualized by politicians. Since simple dictionary methods are not sufficient to fulfil this task, we proceed by developing a framework that allows us to test both hypotheses outlined above.

Sentence Embeddings and the Meaning of Morality

Given the ideological differences in moral foundations among the public, there is good reason to expect similar patterns in political elite communication. However, we saw that politicians frequently use moral language in a way that is inconsistent with the expectations of MFT. At the same time, previous studies are limited in the sense that they only relied on basic dictionary term frequencies to assess moral rhetoric in text. As Haidt (2012) suggested in his statement about fairness, morality can manifest not only in terms of word choice but also in terms of differences in their underlying meaning (see also Chiu et al. 1997). In other words, the same term can take on different meanings depending on the semantic context in which it is used. To our knowledge, Sterling and Jost's (2018) examination of moral discourse on Twitter is the only example that studies these dynamics as part of an exploratory analysis. By extracting individual words that are more likely to appear in the close vicinity of moral terms, the authors find suggestive evidence that politicians rely on similar moral words to emphasize different policy priorities. In the following, we provide a framework to test the same argument more systematically.

A growing body of research in political science started to explore the changing meaning of words across documents – for instance, to track temporal change – by relying on word embeddings (for example, Rheault and Cochrane 2020; Rodman 2020). Put simply, word embeddings represent the vocabulary in a document as vectors in a k-dimensional space where distances between vector representations describe the semantic similarity of the underlying concepts (Rodriguez and Spirling 2022). Recent studies explore the changing meaning of words by examining changes in the underlying embedding vectors of a fixed term (Rodman 2020). However, such an approach requires larger amounts of text to learn embedding vectors than what is available in our context.

We propose an alternative approach that can be applied using much smaller text corpora. Instead of training new word embedding models to compare liberal and conservative politicians, we rely on pre-trained word embeddings (GloVe) to derive embedding representations of the sentences in which moral terms appear. Specifically, we rely on the sentence embedding algorithm proposed by Arora, Liang, and Ma (2016), which conceptualizes the probability of word w being emitted in sentence s as conditional on a *discourse vector* c_s. Crucially, the discourse vector c_s can be understood as representing 'what is being talked about' in sentence s. The assumed data-generating process of any word appearing in a given text then takes the following form:

$$\Pr[w \text{ emitted in sentence } s | c_s] = \alpha p(w) + (1 - \alpha) \frac{\exp(\langle \tilde{c}_s, v_w \rangle)}{Z_{\tilde{c}_s}},$$

$$\text{where } \tilde{c}_s = \beta c_0 + (1 - \beta)c_s, \ c_0 \perp c_s \tag{1}$$

Here, α and β are scalar hyperparameters and $Z_{\tilde{c}_s} = \sum_{w \in v} \exp(\langle \tilde{c}_s, v_w \rangle)$ is a normalizing constant. Importantly, Arora, Liang, and Ma (2016) show that the maximum likelihood estimate of c_s can be approximated by a weighted average of the (pre-trained) embedding vectors of the words in the sentence. In other words, we can use a combination of all word embeddings contained in a sentence to compute a numeric representation of the sentence's underlying content.[1]

---

[1] See Appendix A for a validation of this approach.

Leveraging the approximation of each sentence's discourse vector derived by Arora, Liang, and Ma (2016), we use the following strategy to explore ideological differentiation in the underlying semantic context in which moral terms are mentioned. First, to identify moral language we rely on the updated Moral Foundations Dictionary (MFD2.0) proposed by Frimer et al. (2019; see also Frimer 2020), which builds upon the original MFD developed by Graham, Haidt, and Nosek (2009). Using the dictionary, we extract all sentences that include explicit moral content. Second, we use pre-trained GloVe embeddings (with $k = 100$) and the algorithm proposed by Arora, Liang, and Ma (2016) to compute embeddings for each sentence that contains a moral term. Importantly, the moral term itself is not included when computing the sentence embedding, which ensures that the embeddings only capture the semantic context and are not confounded by the dictionary. Roughly, these embeddings can be understood as a multi-dimensional numeric representation of 'what is being talked about' in each sentence within a given moral foundation. The third step is to investigate the similarity between sentences we have identified within each moral foundation. A standard approach to measure the similarity between embedding vectors is to compute their cosine similarity (for example, Manning et al. 2008). More formally, the cosine similarity is defined as:

$$\text{Cosine similarity} = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \, ||\vec{b}||}, \tag{2}$$

where $\vec{a}$ is the embedding vector of sentence a and $\vec{b}$ is the respective embedding vector of sentence b. In essence, this measure describes the angle between both embedding vectors to quantify their similarity in the embedding space. In practice, the measure ranges from 0 (no semantic overlap) to 1 (perfect semantic overlap).[2] As we will further discuss below, this framework allows us to directly compare differences in the semantic context of the moral language used by political elites, which ultimately enables us to make inferences about the diverging meaning of moral terms (see Rodriguez, Spirling, and Stewart 2023, for similar arguments regarding the relationship between semantic context and the meaning of a word). In contrast to alternative approaches to capture changing meanings of individual terms (for example, Rheault and Cochrane 2020; Rodman 2020), our procedure does not require newly trained embeddings, which makes it ideal for applications with small text corpora. Before discussing these analyses in more detail, however, we first turn to a brief overview of our data sources.

Study 1: How Elites Address National Audiences

We rely on four main data sources from the US and the UK for our analysis of moral rhetoric in elite messaging: (a) State of the Union Addresses given by US presidents (1950–2020), (b) Queen's Speeches in the UK (1950–2017), (c) US convention speeches (1932–2016), and (d) British party leader speeches given at party conferences (1946–2017). The four data sources provide a total of 349 unique documents. Table 1 presents an overview of the different data sources with information on the number of documents and the average word count per speech.

The State of the Union Address and the Queen's Speech receive a lot of attention in the media and, therefore, reach a broad national audience in each respective country. Both speeches are given to inform the legislatures about the government's agenda and, consequently, the speeches cover a wide range of political issues. These speeches are intended for the general public and have been studied for various purposes such as extracting issue salience (Green-Pedersen and Walgrave

---

[2]Note that cosine similarities are not strictly bound to be positive since the angle between two vectors can exceed $90°$. Thus, while cosine similarities of embedding vectors usually remain in positive space, the measure can take any value in the interval $[-1, 1]$.

Table 1. Overview of speeches

Speech Type	Number of Documents	Average Word Count
US State of the Union	80	6,780
UK Queen's Speeches	65	1,096
US Convention Speeches	39	4,158
UK Party Leader Speeches	165	5,870

2014; Hobolt and Klemmensen 2008), ideological positions (Hakhverdian 2010), and emotive rhetoric (Kosmidis et al. 2019).

In comparison, the speeches given by presidential candidates at party conventions in the US and the party leader speeches given at yearly conferences in the UK are delivered in front of in-party supporters. While these speeches still target national audiences, the messaging is focused more on party members in an attempt to mobilize the politician's base for upcoming elections. Party leader and convention speeches are, therefore, not as encompassing as the State of the Union Address or the Queen's Speech. Accordingly, this comparison provides an initial opportunity to examine whether politicians adjust their moral rhetoric to appeal to different audiences.

By combining speeches from the US and the UK, we are further able to compare ideological differences in moral rhetoric across political systems. Despite the presence of a socialist party in the UK and the more important role of religion in US politics, both political systems are dominated by two major political parties and have relatively similar political cultures. Thus, we can directly assess whether our findings replicate across similar political environments.

Liberal and Conservative Politicians use the Same Moral Terms

Do politicians across the ideological spectrum use moral rhetoric consistent with the predictions of MFT? Fig. 1 compares party differences in the use of moral language across the four sets of speeches using the updated Moral Foundations Dictionary (MFD2.0) proposed by Frimer et al. (2019). Specifically, the figure displays the average percentage of dictionary terms related to each moral dimension that appears in a given speech. Remarkably, the overall emphasis on moral foundations is quite consistent across speeches, even though we are comparing contexts with varying target audiences. Overall, politicians tend to use more care-, loyalty-, and authority-related words while language about fairness or sanctity appears less common.

More importantly, however, we do not find systematic ideological differences that match theoretical expectations in the literature. Across data sources, there is little evidence for diverging moral rhetoric and to the extent that significant differences between liberal and conservative politicians exist they are not necessarily consistent with the expectations of MFT. Taking the US State of the Union Address as an example, Democratic presidents emphasize the loyalty dimension slightly more than their Republican counterparts.[3] There is no evidence that differential moral rhetoric only materializes for more targeted audiences (that is, in UK party leader speeches or US convention speeches). In sum, we are unable to detect the clear systematic differences between the major parties on the five dimensions of morality predicted by MFT.

A possible explanation for the lack of ideological differences could be that the comparison between parties obscures systematic variation in moral language over time. For example, historic events or gradual policy shifts might shape when parties use different moral terms. Therefore, in Appendix B, we examine how the emphasis on specific moral foundations changed over time for each speech type. However, there is no clear evidence that the use of moral language over time

---

[3]While this difference does not reach conventional levels of statistical significance, it is worth noting that Sterling and Jost (2018) similarly found that liberal members of Congress used more language pertaining to group loyalty than conservative lawmakers on Twitter.
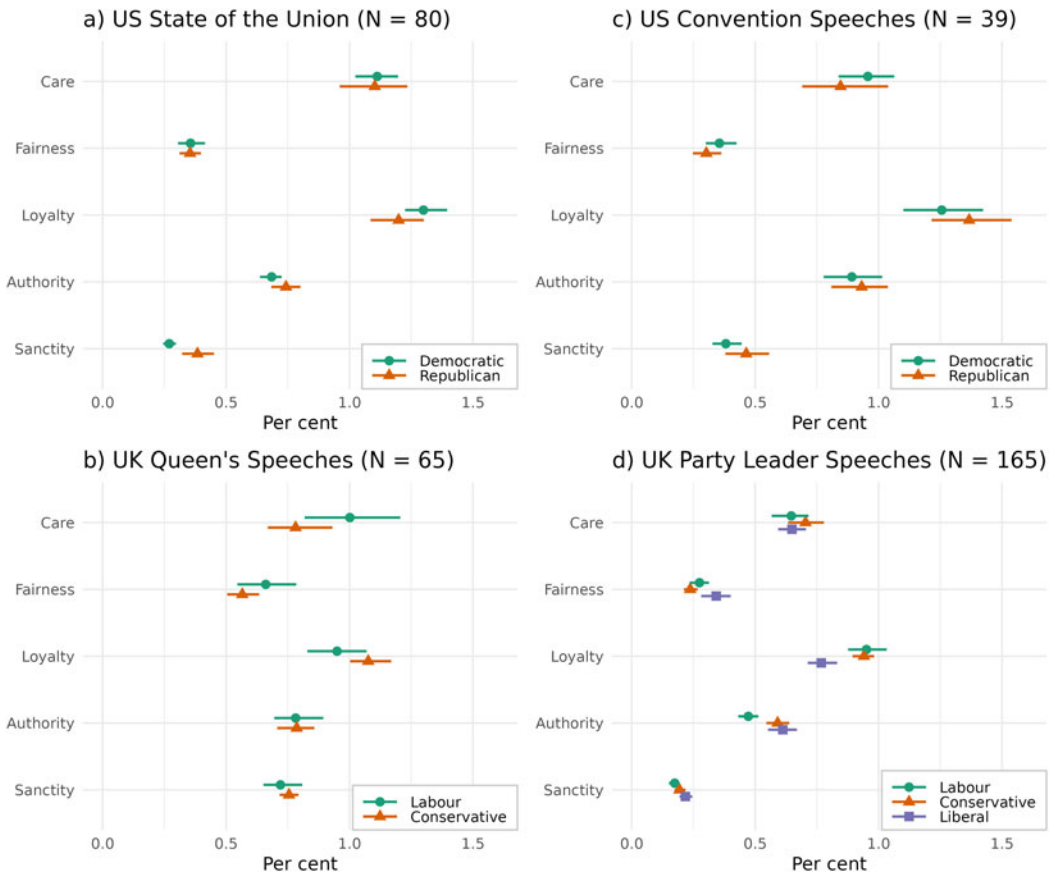
Figure 1. Average mention of moral dimensions in all data sources: 95 per cent confidence intervals based on a non-parametric document-level bootstrap.

reflects gradual policy shifts or historic events that causes potential discontinuities. As such, the average proportions covering the entire periods under consideration do not conceal theoretically meaningful variation in moral foundations.

Another explanation could be that there are substantial ideological differences in moral language use within rather than between each foundation. In other words, the categories in the MFD might be too broad, such that liberals and conservatives appear to talk about the same underlying foundation but emphasize systematically different words. To rule out this possibility, we examine whether politicians also tend to rely on the same terms within each moral dimension. Appendix C shows that the most common moral words are used in similar amounts by liberal and conservative politicians, which indicates that there is not only a lack of systematic differentiation in moral foundations as a whole but also in the use of individual dictionary terms themselves.

Accordingly, political elites not only emphasize each moral foundation to a similar degree but also use the same terms within each dimension. This lack of ideological differentiation in moral messaging is consistent across speeches that target national audiences as well as party supporters, a finding that appears to contradict *H1*. However, as we suggest in *H2*, even if politicians across the political spectrum use the same moral terms within each foundation, this does not necessarily imply that they attach the same underlying meaning to these terms. We investigate this possibility in the remainder of this article.

### Moral Foundations Can Have Diverging Meanings

To explore whether elites with different ideological standings use the same moral terms in diverging semantic contexts, we use the following procedure for each speech type and moral foundation. First, we select all sentences containing at least one dictionary term belonging to the corresponding foundation. Next, we compute the embedding cosine similarities between all cross-document sentence pairs. Recall that this cosine similarity can be understood as a measure of the semantic overlap of 'what is being talked about' in each sentence. Next, iterating over individual documents, we compute the average difference in cosine similarities of statements by politicians belonging to the same party versus those belonging to a different party.[4] This provides us with a single numeric summary for each speech, indicating whether the semantic context of moral terms belonging to each foundation is more coherent with other in-party or out-party speeches. A positive value on this metric implies that, within a given moral foundation, on average, sentence pairs from two different speeches made by politicians from the same party are more similar than sentence pairs belonging to different parties. Note that this approach directly accounts for within-party variation in the context of moral language use. We re-scale the measure to standard deviation units to facilitate comparisons between speech types and compute uncertainty estimates using a non-parametric document-level bootstrap.

Using this method, we compute average differences in cosine similarities for each moral dimension within each speech type. The results are displayed in Fig. 2. To reiterate, positive values indicate that the semantic contexts of a given moral foundation – in other words, approximations of 'what is being talked about' – are more similar within parties than they are between parties. While our previous analysis indicates that there were no systematic differences between liberal and conservative politicians' emphasis on distinct moral dimensions, we now find evidence for diverging moral rhetoric on certain foundations.

In the US, the semantic context of moral foundations in political speeches remained largely consistent across party lines. Only politicians' discussion of loyalty suggests a certain degree of ideological differentiation of moral rhetoric in State of the Union Addresses as well as party convention speeches (both $p < 0.10$). In other words, Republican and Democratic presidents and presidential candidates used moral terms centred around loyalty in different semantic contexts and, therefore, attached diverging meanings to the foundation. This finding is particularly interesting since Democratic presidents emphasized the loyalty foundation slightly more than Republicans in State of the Union Addresses (see Fig. 1), a pattern that is not consistent with the predictions of MFT. Our analysis of the semantic context of moral terms helps contextualize this seeming incongruity between theory and empirics: Democratic politicians may emphasize loyalty more than Republicans, but they essentially mean different things when talking about this foundation.

In the United Kingdom, on the other hand, the moral foundation with the clearest divergence in the semantic context is fairness rather than loyalty. Thus, while the competition over the meaning of morality is mostly limited to the loyalty found in the US, it appears more comprehensive and particularly pronounced for the fairness foundation in the UK. Again, this analysis helps us contextualize our initial evidence focused on the raw proportions of moral foundation terms. Although there is some indication that Labour politicians emphasized fairness more than Conservative politicians in the Queen's Speeches and party leader speeches, these differences do not reach conventional levels of statistical significance (see Fig. 1). However, taking into account the semantic context shows that there is clear ideological differentiation in moral rhetoric between liberal and conservative elites on this dimension. In sum, these results suggest that, in the UK, moral rhetoric surrounding the fairness foundation is especially contested even though politicians across the political spectrum emphasize the foundation to a similar degree.

---

[4]The cosine similarities are only based on inter-document sentence comparisons. In other words, the within-party similarity measure is not confounded by larger semantic overlap within a given document itself.
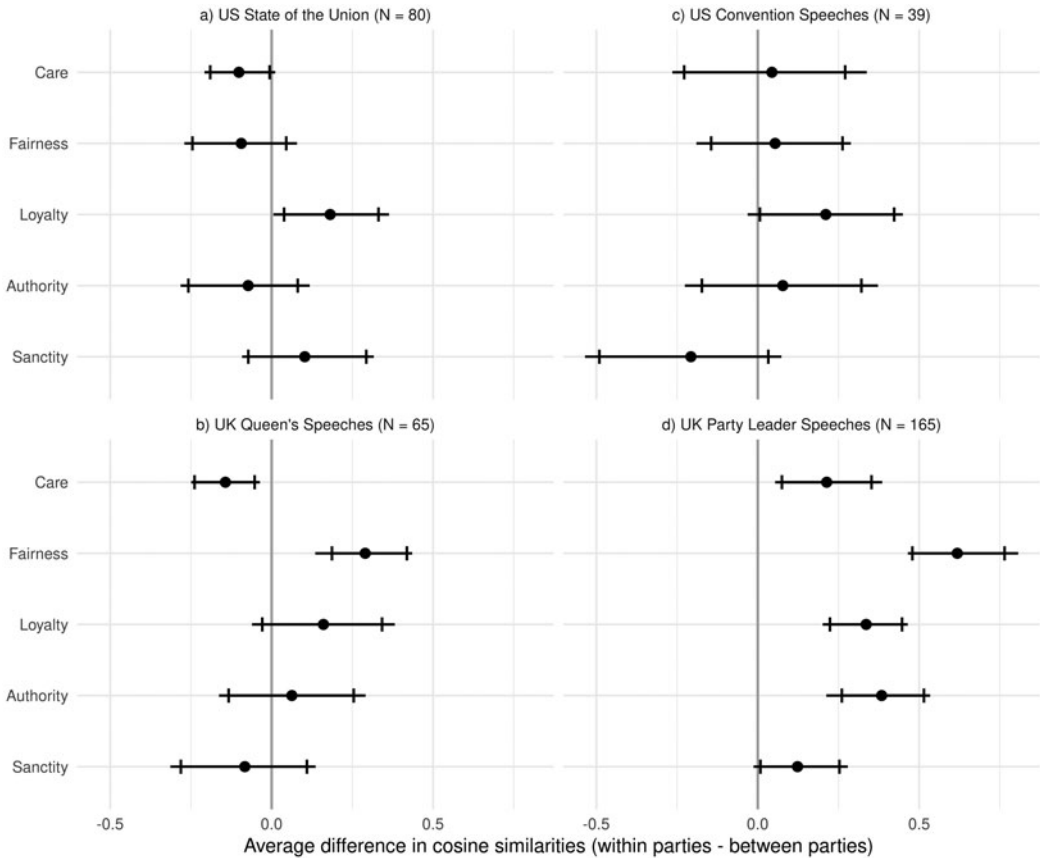
Figure 2. The average difference in cosine similarities between embedding vectors of sentences mentioning each moral foundation in standard deviations (within parties – between parties). Positive values indicate that sentences mentioning a given foundation are more similar within the same party than between parties: 95 per cent and 90 per cent confidence intervals based on a non-parametric document-level bootstrap.

Comparing the results between speech types, we see that ideological differentiation across all foundations is stronger in party leader speeches than in Queen's Speeches. The significant difference in cosine similarities is not only larger on the fairness foundation, but also manifests in the care, loyalty, and authority dimensions. Finding stronger ideological differentiation in party leader speeches lends support to *H1* since these speeches address more narrow partisan audiences than the Queen's Speeches. Lastly, it is noteworthy that when looking at the Queen's Speeches, the semantic context of moral terms belonging to the care foundation varies more within parties than between parties, as indicated by the negative average difference in cosine similarities. This result may suggest that when it comes to care, moral divergence is better characterized as changing rhetoric across time rather than between parties.

While the results thus far indicate that ideological differentiation in the semantic contexts of moral foundations is stronger in the UK than in the US, it is worth noting that this may not be the case in the future. Figure 3 explores average differences in cosine similarities over time. This comparison allows us to examine whether moral rhetoric has become increasingly differentiated – for example, due to partisan polarization. Indeed, in the US there is evidence that ideological differentiation in moral rhetoric is increasing over time for both types of speeches. To the extent that this trend continues, we expect that the partisan divergence in moral rhetoric in the US will
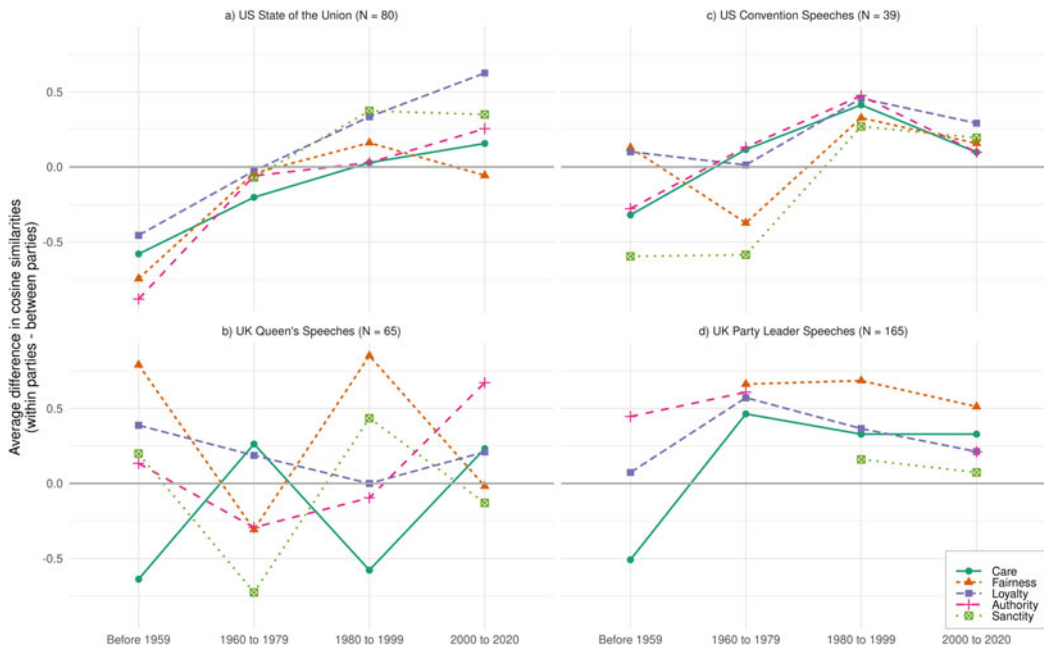
Figure 3. Average differences in cosine similarities between embedding vectors of sentences mentioning each moral foundation in standard deviations (within parties – between parties) over time. Positive values indicate that sentences mentioning a given foundation are more similar within the same party than between parties.

become more pronounced in the upcoming years. There is no comparable trend towards more ideological differentiation in the United Kingdom. If anything, average differences in cosine similarities appear to decline when looking at party leader speeches. Thus, while divergence in the meanings of morality was most evident in UK party leader speeches, the trends suggest that moral discourse in the US may soon show similar patterns.

In sum, our analyses illustrate the great potential of moving beyond basic dictionary approaches to study the politics of morality in elite messaging. Although politicians across the ideological spectrum might use the same moral terms, they appear to attach diverging meanings to them. Importantly, to capture these nuances in moral divergence, scholars have to supplement dictionary counts by incorporating the semantic context in which moral terms appear. For instance, our analysis thus far illustrates that some of the findings that initially appear inconsistent with MFT may be explained by the fact that the underlying dictionary terms differ in meaning between liberal and conservative politicians.

Study 2: Moral Divergence Across Contexts

Analyzing four types of speeches in two countries, we found evidence for ideological differentiation in moral rhetoric, particularly about the dimensions of fairness and loyalty. Moral divergence appears more pronounced when political elites address their respective supporters (for example, at party conventions) rather than broader audiences (for example, at the State of the Union Address). That said, we have to consider two potential caveats when interpreting these results. First, our measure of ideological differentiation in the meanings of morality may be confounded by the distribution of topics across speeches. After all, a complete lack of overlap in topics covered by liberal and conservative speeches, albeit unlikely, would make it challenging attribute differences in semantic context solely to diverging meanings of morality. Second,

although party leader and convention speeches are mostly targeted at partisan supporters, they nevertheless receive national media attention and may, therefore, still reach a fairly broad audience. This raises the question of how moral divergence manifests in a context where messages can be disseminated exclusively to a narrower audience. We address both caveats in a second set of analyses, leveraging additional text corpora from the US: that is, transcripts of presidential debates and the content of congressional newsletters sent by senators to their supporters via email.

Moral Divergence Persists When Elites Discuss the Same Topics

The first concern raised above is the potential confounding due to document-level variation in topics discussed by politicians. We evaluate this alternative explanation by replicating our analysis using a text corpus that represents a least-likely case for moral divergence if our findings were solely driven by the sets of topics raised in speeches. Specifically, we compare ideological differences in moral rhetoric in transcripts of presidential and vice-presidential debates in the US covering the years from 1960 to 2016.

Using debate transcripts provides two crucial advantages to supplement our previous analysis. First, we can explore ideological differentiation in moral rhetoric when elite messaging emerges spontaneously instead of being carefully scripted. Unlike in speeches, participants in the presidential and vice presidential debates use moral terms in a setting where they do not have the opportunity to carefully prepare every statement they plan to deliver, especially when it comes to defending their position in a back-and-forth argument. Since debates are dynamic by nature, participants have to provide answers to unanticipated questions and attacks, which enables us to investigate moral language that emerges more naturally than in a prepared speech. Secondly, the debate environment ensures that both contestants discuss the same underlying topics. The fact that the debates are moderated and candidates answer the same set of questions guarantees that the broader set of topics being covered is held constant between liberal and conservative candidates. In short, the presidential debates provide a useful supplemental test to investigate ideological differentiation in moral rhetoric when political elites have no opportunity to carefully prepare their statements and have no direct control over the set of topics being covered.

We use the same procedures as in Study 1 to extract the overall proportion of moral foundation terms as well as differences in the semantic context of each dimension. The results are summarized in Fig. 4. Even though by definition, the debate participants are addressing the same general audience, there are more distinct ideological differences in the relative emphasis on
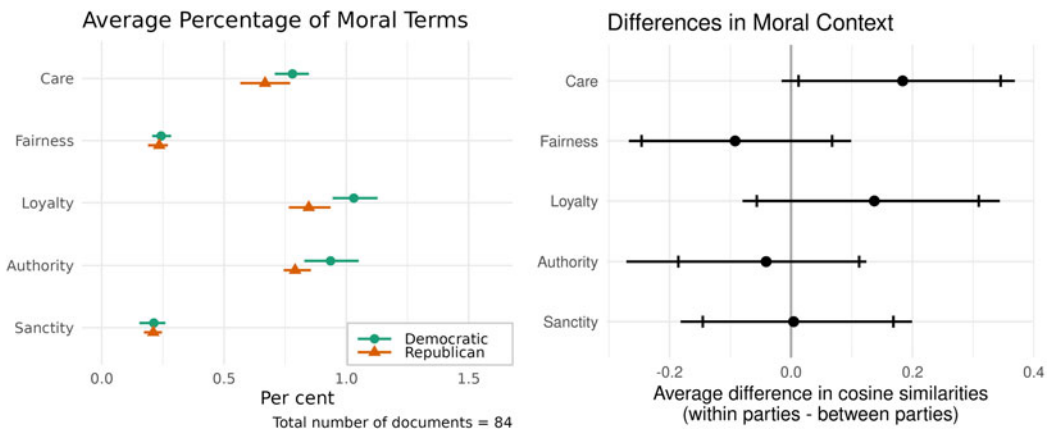
Figure 4. Replication of Study 1 using transcripts from US presidential and vice-presidential debates (1960–2016): 95 per cent and 90 per cent confidence intervals based on non-parametric document-level bootstrap.

individual moral foundations than in State of the Union Addresses or party convention speeches. Specifically, Democratic candidates were more likely than Republicans to mention terms belonging to the care ($p < 0.10$), loyalty ($p < 0.05$), and authority ($p < 0.10$) foundations, the latter two differences standing in contrast to predictions derived from MFT. This rising ideological discrepancy in the relative emphasis on moral foundations may be because debate utterances are less scripted than written speeches, especially in situations when the exchange between candidates is becoming argumentative. Thus, candidates may be less able to maintain ambiguous moral stances in spontaneous statements even though they address the same national audience.

What about ideological differentiation in terms of the semantic context of moral terms? Again, comparing cosine similarities of sentences mentioning each moral foundation helps us contextualize the differences in moral rhetoric. First, when it comes to the care foundation, Democratic and Republican candidates not only differ in average mentions, but they also use the terms in different semantic contexts ($p < 0.10$). Secondly, regarding the loyalty foundation – where we found the clearest evidence of ideological differences that contradicted MFT – there is only weak and statistically insignificant evidence for divergence in semantic contexts. This pattern is even stronger for the authority dimension where we find practically no evidence that the semantic context of dictionary terms differs between Republican and Democratic candidates. Thus, in contrast to the theoretical expectations of MFT, Democratic candidates in presidential debates mention more terms belonging to the loyalty and authority dimensions than Republicans and, in addition, these terms appear in the same contexts across parties. A potential interpretation of this finding may be that Democratic candidates make concerted attempts to appeal to conservative voters in these debates. Another explanation for the unexpected ideological differentiation on the loyalty foundation may be that Democratic candidates have to foster unity within more diverse coalitions of supporters than Republicans (see Grossmann and Hopkins 2016).

To provide more intuition about the nature of the differences in semantic contexts in which moral terms appear, we present a sample statement by Hillary Clinton in one of the 2016 presidential debates in Table 2. This sentence mentions the term 'share', which belongs to the care foundation in the MFD. In it, Clinton argues for closing tax loopholes and having the wealthy pay their fair share. Out of all remaining statements in the 2016 presidential debates included in our corpus, we extracted all sentences that also mention the term 'share' and selected the two most similar and most dissimilar sentences according to their cosine similarity with Clinton's original statement. Interestingly, they all discuss the necessity to pay a 'fair share', albeit in very different contexts. The two most similar sentences discuss tax reform using almost the same wording as Clinton's initial statement. Notably, both statements were made by Clinton rather than Trump. The two most dissimilar sentences, on the other hand, raise the issue of paying a fair share in a strikingly different context, namely the defence budgets of NATO countries.

Table 2. Sample statements made by Hillary Clinton during the presidential debates that mention the term 'share' along with other debate statements that mention the same term and are most similar/dissimilar to Clinton's statement according to the cosine similarity of sentence embeddings

Description	Speaker	Year	Sentence	Similarity
Original	Clinton	2016	we're going to do it by having the wealthy pay their fair share and close the corporate loopholes.	1.00
Similar	Clinton	2016	and I think it's time that the wealthy and corporations paid their fair share to support this country.	0.90
	Clinton	2016	so we are going to have the wealthy pay their fair share.	0.83
Dissimilar	Trump	2016	number one, the countries of NATO, many of them aren't paying their fair share.	0.54
	Trump	2016	all I said, that it's very possible that if they don't pay a fair share, because this isn't years ago where we could do what we're doing.	0.47

Here, both statements were made by Donald Trump, which exemplifies the pattern of larger semantic congruence within parties than between parties.[5]

This comparison illustrates how our sentence embedding approach captures whether the same moral terms are used in different contexts, and it also shows how these contexts are inherently intertwined with variations in topics.[6] Although we focused on presidential debates where both candidates have to discuss the same topics throughout the conversation, our examples of dissimilar statements ultimately reveal how the same moral term is raised in the context of different political issues. In other words, while the general distribution of topics within a debate should be similar between Democratic and Republican candidates, they still deviate concerning the context in which the same moral terms appear.

Consistent with other approaches that rely on text embeddings, we understand these contextual differences as a manifestation of divergent 'meanings' of the underlying terms (Rodman 2020; Rodriguez, Spirling, and Stewart 2023). This interpretation rests on the distributional hypothesis in linguistic theory, which posits that words that appear in the same context ultimately have similar meanings (Firth 1957; Harris 1954). Furthermore, our analysis of presidential debates indicates that these contextual differences are not just driven by document-level variation in topic proportions, but rather by diverging semantic contexts within documents. Ultimately, the finding that liberal and conservative politicians talk about care or fairness in different contexts suggests that they are likely to have a diverging understanding of the underlying moral connotation.

In sum, there is substantial variation, both between and within parties, in terms of how particular moral terms are used. While more work is needed to conceptually disentangle variations in topics and the underlying meaning of morality, our sentence embedding framework provides a first step for such a more nuanced analysis of ideological differentiation in moral rhetoric.

Exclusive Messages Exacerbate Moral Divergence

Our second concern pertains to the question of whether politicians have sufficient incentives for moral divergence in the speeches considered thus far. Although we explored variations in audience composition (in-party supporters vs. broader population), all of the speeches received substantial media attention and were, in principle, accessible to the entire electorate. To address this potential shortcoming, we leveraged an additional set of documents, consisting of direct messages from politicians to their constituents. More specifically, we examined a random sample of 1,000 newsletter emails that were sent to the supporters of US senators in 2020.[7]

After collecting the sample of emails, we again used the same procedures as in Study 1 to extract the overall proportion of moral foundation terms as well as differences in the semantic context of each dimension. The results are displayed in Fig. 5. In a context where politicians speak directly and exclusively to their supporters, moral divergence is substantially more pronounced in terms of both the relative emphasis on individual moral foundations as well as the semantic context of moral terms. Similar to previous results, however, differences between liberal and conservative politicians are not always consistent with the predictions of MFT.

Figure 4a reveals that Democratic senators emphasize the care, loyalty, and sanctity foundation more than their Republican counterparts. Democratic senators' higher emphasis on loyalty and sanctity, as well as the lack of party differences on the fairness dimension, directly contradict MFT. As before, our embedding analysis presented in Fig. 4b helps to contextualize these findings. Take fairness for example: based on the mentioning of the moral foundation alone, there

---

[5]Note that the comparison in Table 2 only serves illustrative purposes. When computing the average differences in cosine similarities reported above, we omit all within-document comparisons.

[6]We thank an anonymous reviewer for raising this important point.

[7]These data come from Lindsey Cormack's (2017) data set of constituent e-newsletters. The full set of emails is accessible via https://www.dcinbox.com/.
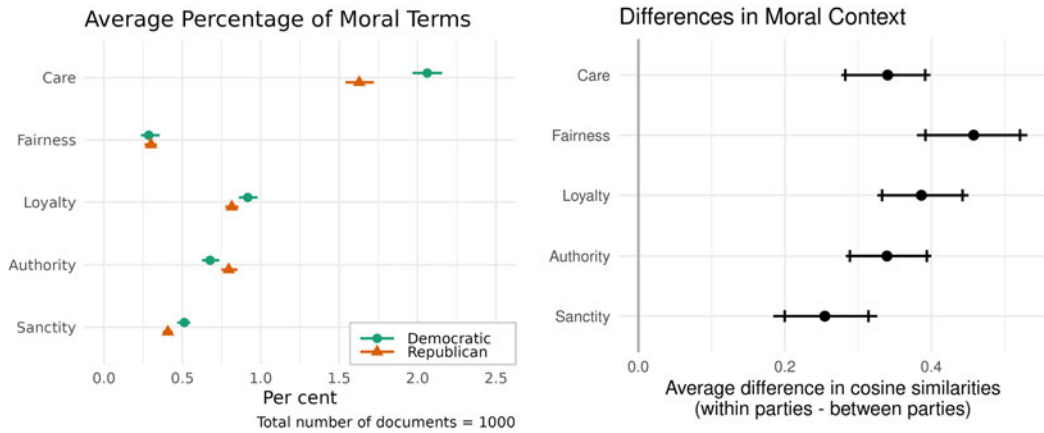
Figure 5. Replication of Study 1 using a sample of 1,000 newsletter emails sent to the supporters of US senators in 2020: 95 per cent and 90 per cent confidence intervals based on non-parametric document-level bootstrap.

appears to be no significant difference between Democratic and Republican senators, again a finding that is inconsistent with MFT. At the same time, however, moral divergence based on the semantic context in which dictionary terms appears is higher for the fairness dimension than any other moral foundation. Although liberal and conservative congressmen appear to emphasize fairness to the same degree, they differ substantially in the underlying meaning they attach to the moral dimension. Similar arguments can be made for the loyalty and sanctity foundation, which appears to be more strongly emphasized by Democratic than Republican senators. To understand ideological differentiation in moral reasoning, what matters is not only the relative degree to which individual foundations are promoted in raw numbers but, rather, how they are contextualized and, ultimately, their underlying meaning.

It is worth emphasizing in this context that the average differences in cosine similarities (that is, our measure of moral divergence) is statistically significant across all five moral foundations. In contrast, in our previous analyses, these differences were only distinguishable from zero for a few individual foundations. One explanation for this finding, of course, is the fact that the number of documents is substantially larger, which implies higher statistical power. At the same time, we maintain there is good reason to believe that this result is, at least partly, driven by differences in audience composition. First, the emails were directly sent to supporters who signed up for the newsletters, which implies that senators can directly speak to a narrow and exclusive audience. Following the theoretical arguments outlined at the beginning of this article, this represents the most likely case for moral divergence in terms of both the relative emphasis of specific foundations as well as their diverging meaning. Furthermore, the emails selected for this analysis were all sent in 2020, whereas previous speech corpora covered periods of up to eighty-four years. Since the meanings of morality change over time and our measure of moral divergence takes into account within-party variation in semantic contexts, it is likely that larger periods increase within-party variation, therefore reducing the relative between-party moral divergence.

Discussion

We examined elite moral rhetoric across a variety of contexts. When it comes to the relative emphasis on moral foundations, ideological differentiation is not as clear as MFT would predict. Some of our findings directly contradict the theory. For instance, Democrats in the US sometimes emphasize the loyalty foundation more than their Republican counterparts. Although similar results have been reported in previous research (for example, Sterling and Jost 2018), these

apparent inconsistencies with MFT exemplify the advantage of our sentence embedding approach since it helps us contextualize observed differences in moral rhetoric (or the lack thereof). In the case of the Democrats' emphasis on the loyalty foundation, for example, we found evidence that these moral terms appear in systematically different contexts than when mentioned by Republicans. A potential explanation for this moral divergence on the loyalty foundation could be the fact that Democrats have to rely on more diverse coalitions of voters than Republicans (see Grossmann and Hopkins 2016). Thus, their campaign rhetoric may require more intense (and different) appeals to loyalty to cultivate greater unity among their supporters.

On the other hand, the presidential debates also revealed a stronger emphasis on the authority dimension among Democratic candidates. In contrast to the loyalty foundation, we found no evidence that these moral terms appear in diverging contexts. This result could either imply a more unequivocal deviation from MFT or it may suggest that Democratic candidates are making systematic attempts to appeal to conservative voters. While more work is necessary to disentangle both potential explanations, our sentence embedding approach provides a first step to differentiate between apparent inconsistencies with MFT and other cases where moral dimensions are simply interpreted differently across the ideological spectrum.

Another possible explanation for the at times inconsistent findings is the need to revise the MFD and its underlying theoretical framework. For instance, previous scholars have proposed the inclusion of a sixth moral foundation, liberty, particularly because it helps differentiate the moral profile of libertarians (for example, Iyer et al. 2012). While an expansion of the MFD is beyond the scope of this article, we believe that these are important and promising areas for future research (see for example Araque, Gatti, and Kalimeri 2021; Rezapour, Shah, and Diesner 2019).

What we are left with is the question of why moral divergence appears so much more pronounced in the case of UK speeches when compared to the US (at least in the context of Study 1). Our theoretical argument surrounding audience composition suggests that UK party leader speeches, where we found the strongest evidence for moral divergence on multiple foundations, may gain less widespread attention than, say, party convention speeches in the US. After all, party conventions in the US are only held in the context of presidential elections and are scrutinized by the liberal and conservative media alike. UK party leader speeches, on the other hand, are held once a year and receive comparatively less media attention. Thus, party leaders in the UK can be presumed to address mostly supporters, which results in stronger moral divergence. This argument is supported by our supplemental analysis in Study 2, which revealed even more pronounced moral divergence in the context of email newsletters sent directly by US senators to their respective supporters. In sum, what appears to matter most in determining moral rhetoric is audience composition and whether their messaging can exclusively target supporters.

At the same time, we have to consider the possibility that our somewhat ambiguous findings on moral divergence may be a consequence of our embedding approach itself. To assess this option, we replicate our analyses using a related alternative approach. In a recent paper, Rodriguez, Spirling, and Stewart (2023) propose a new method to compare the meaning of individual words across different sets of documents based on pre-trained word embeddings. Instead of computing individual sentence embeddings for each occurrence of a term included in the dictionary (as in our approach), they compute average embeddings of a target term itself across each set of documents. We implement their embedding regression framework for each moral foundation and compare the resulting embedding vectors between liberal and conservative politicians (see Appendix D). In contrast to our findings, the replication using the method proposed by Rodriguez, Spirling, and Stewart (2023) suggests that all moral foundations across all types of documents are significantly different between liberal and conservative politicians. Crucially, however, their approach fails to account for the within-party variation in a semantic context. Thus, Rodriguez, Spirling, and Stewart (2023) overestimate the extent to which context embeddings differ between liberal and conservative politicians relative to the varying semantic context within each group.

Conclusion

While the fundamental question of what is morally right or wrong is central to politics, exploring the underlying dynamics of moral rhetoric remains an empirical challenge. Here, using the state-of-the-art dictionary to measure moral foundations in text, we began our study by examining whether and how moral rhetoric differs between liberal and conservative political elites. In contrast to theoretical expectations established in the literature, we found no consistent ideological differences in the average emphasis on moral foundations across five types of political speeches and debates from two countries covering a period of more than fifty years. Despite the previous success of MFT at establishing core relationships between the five dimensions of morality and ideology in the mass public, we find no evidence for equivalent patterns in political elite rhetoric.

Compared to previous studies in this area, however, we move beyond simple dictionary methods by employing a novel sentence embedding approach, allowing us to directly compare the semantic context in which liberal and conservative politicians use the same moral terms. In sum, the results show that politicians rely on moral language in more nuanced ways than hitherto assumed in the literature. In the US, we find ideological differences in the semantic context of moral language use focused on the loyalty dimension. Evidence from the UK, on the other hand, suggests that elite rhetoric is particularly contested surrounding the fairness dimension. Overall, even though politicians across the ideological spectrum tend to rely on the same moral terms in their speeches, they appear to use them in systematically different contexts and attach diverging meanings to them. This moral divergence, in turn, is particularly pronounced in contexts where elites can directly address a narrow and exclusive audience.

Together, we view our results as a first step towards a better understanding of the role of morality in political communication, and there are fruitful avenues for future research to disentangle when and how particular moral values are evoked in a given context. Overall, we recommend caution when applying standard dictionary approaches alone to capture ideological differences in moral rhetoric since these fail to account for the semantic context in which politicians use moral language. From a more general methodological perspective, the sentence embedding approach developed here provides a useful tool to validate other dictionaries, especially in contexts where researchers are worried about similar terms being used in diverging contexts.

References

Araque O, Gatti L and Kalimeri K (2021) The Language of Liberty: A preliminary study. In *Companion Proceedings of the Web Conference 2021*, pp. 623–26.

Arora S, Liang Y and Ma T (2016) A simple but tough-to-beat baseline for sentence embeddings. In *International Conference on Learning Representations* (Toulon, France), pp. 1–16.

Barker DC and Tinnick JD (2006) Competing visions of parental roles and ideological constraint. *American Political Science Review* **100**(2), 249–63.

Blumenau J and Lauderdale BE (2022) The variable persuasiveness of political rhetoric. *American Journal of Political Science* forthcoming, 1–16.

Bos L and Minihold S (2022) The ideological predictors of moral appeals by European political elites; An exploration of the use of moral rhetoric in multiparty systems. *Political Psychology* **43**(1), 45–63.

Brady WJ et al. (2019) An ideological asymmetry in the diffusion of moralized content on social media among political leaders. *Journal of Experimental Psychology: General* **148**(10), 1802.

Chiu C-Y et al. (1997) Implicit theories and conceptions of morality. *Journal of Personality and Social Psychology* **73**(5), 923–40.

Clifford S (2017) Individual differences in group loyalty predict partisan strength. *Political Behavior* **39**(3), 531–52.

Clifford S and Jerit J (2013) How words do the work of politics: Moral foundations theory and the debate over stem cell research. *The Journal of Politics* **75**(3), 659–71.

Clifford S et al. (2015) Moral concerns and policy attitudes: Investigating the influence of elite rhetoric. *Political Communication* **32**(2), 229–48.

Cormack L (2017) DCinbox: Capturing every congressional constituent e-newsletter from 2009 onwards. *The Legislative Scholar* **2**(1), 27–34.

Day MV et al. (2014) Shifting liberal and conservative attitudes using moral foundations theory. *Personality and Social Psychology Bulletin* **40**(12), 1559–73.

Downs A (1957) *An Economic Theory of Democracy*. New York: Harper and Row.

Easton D (1953) *The Political System*. New York: Alfred A. Knopf.

Enke B (2020) Moral values and voting. *Journal of Political Economy* **128**(10), 3679–3729.

Feinberg M and Willer R (2013) The moral roots of environmental attitudes. *Psychological Science* **24**(1), 56–62.

Feinberg M and Willer R (2015) From gulf to bridge: When do moral arguments facilitate political influence? *Personality and Social Psychology Bulletin* **41**(12), 1665–81.

Firth JR (1957) *Studies in Linguistic Analysis*. Oxford: Wiley-Blackwell.

Frimer JA (2020) Do liberals and conservatives use different moral languages? Two replications and six extensions of Graham, Haidt, and Nosek's (2009) moral text analysis. *Journal of Research in Personality* **84**, 103906.

Frimer JA et al. (2019) Moral foundations dictionary for linguistic analyses 2.0. (*Unpublished manuscript*)

Graham J, Haidt J and Nosek BA (2009) Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology* **96**(5), 1029–46.

Graham J et al. (2013) Moral foundations theory: The pragmatic validity of moral pluralism. In Devine P and Plant A (eds), *Advances in Experimental Social Psychology*, vol. 47. Cambridge: Academic Press, pp. 55–130.

Green-Pedersen C and Walgrave S (2014) Political agenda setting: An approach to studying political systems. In Green-Pedersen C and Walgrave S (eds), *Agenda Setting, Policies, and Political Systems*. Chicago: The University of Chicago Press, pp. 1–18.

Grossmann M and Hopkins DA (2016) *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. New York: Oxford University Press.

Haidt J (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review* **108**(4), 814.

Haidt J (2012) *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Vintage.

Haidt J and Joseph C (2004) Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus* **133**(4), 55–66.

Haidt J, Koller SH and Dias MG (1993) Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology* **65**(4), 613.

Haidt J, Graham J and Joseph C (2009) Above and below left–right: Ideological narratives and moral foundations. *Psychological Inquiry* **20**(2–3), 110–19.

Hakhverdian A (2010) Political representation and its mechanisms: A dynamic left–right approach for the United Kingdom, 1976–2006. *British Journal of Political Science* **40**(4), 835–56.

Harris ZS (1954) Distributional structure. *Word* **10**(2–3), 146–62.

Hatemi PK, Crabtree C and Smith KB (2019) Ideology justifies morality: Political beliefs predict moral foundations. *American Journal of Political Science* **63**(4), 788–806.

Hobolt SB and Klemmensen R (2008) Government responsiveness and political competition in comparative perspective. *Comparative Political Studies* **41**(3), 309–37.

Hochschild JL (1981) *What's Fair? American Beliefs About Distributive Justice*. Cambridge: Harvard University Press.

Iyer R et al. (2012) Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PloS One* **7**(8), e42366.

Jung J-H (2020) The mobilizing effect of parties' moral rhetoric. *American Journal of Political Science* **64**(2), 341–55.

Kertzer JD et al. (2014) Moral support: How moral values shape foreign policy attitudes. *Journal of Politics* **76**(3), 825–40.

Kim KR, Kang J-S and Yun S (2012) Moral intuitions and political orientation: Similarities and differences between South Korea and the United States. *Psychological Reports* **111**(1), 173–85.

Koleva SP et al. (2012) Tracing the threads: How five moral concerns (Especially Purity) help explain culture war attitudes. *Journal of Research in Personality* **46**(2), 184–94.

Kosmidis S et al. (2019) Party competition and emotive rhetoric. *Comparative Political Studies* **52**(6), 811–37.

Kraft PW (2018) Measuring morality in political attitude expression. *The Journal of Politics* **80**(3), 1028–33.

Kraft PW and Klemmensen R (2023) Replication Data for: Lexical Ambiguity in Political Rhetoric: Why Morality Doesn't Fit in a Bag of Words. Available at https://doi.org/10.7910/DVN/GNIJVE, Harvard Dataverse, V1.

Lakoff G (2010) *Moral Politics: How Liberals and Conservatives Think*. Chicago: The University of Chicago Press.

Lipsitz K (2018) Playing with emotions: The effect of moral appeals in elite rhetoric. *Political Behavior* **40**(1), 57–78.

Low M and Wui MGL (2015) Moral foundations and attitudes towards the poor. *Current Psychology* **35**, 650–56.

Manning CD, Raghavan P, and Schütze H (2008) *Introduction to Information Retrieval*. Cambridge: Cambridge University Press

Meegan D (2019) *America the Fair: Using Brain Science to Create A More Just Nation*. Ithaca: Cornell University Press.

Nilsson A and Erlandsson A (2015) The moral foundations taxonomy: Structural validity and relation to political ideology in Sweden. *Personality and Individual Differences* **76**, 28–32.

Page BI (1976) The theory of political ambiguity. *American Political Science Review* **70**(3), 742–52.

Rezapour R, Shah SH and Diesner J (2019) Enhancing the Measurement of Social Effects by Capturing Morality. In *Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pp. 35–45.

Rheault L and Cochrane C (2020) Word embeddings for the analysis of ideological placement in parliamentary corpora. *Political Analysis* **28**(1), 112–33.

Rodman E (2020) A timely intervention: Tracking the changing meanings of political concepts with word vectors. *Political Analysis* **28**(1), 87–111.

Rodriguez PL and Spirling A (2022) Word embeddings: What works, what doesn't, and how to tell the difference for applied research. *The Journal of Politics* **84**(1), 101–15.

Rodriguez PL, Spirling A and Stewart BM (2023) Embedding regression: Models for context-specific description and inference. *American Political Science Review* **forthcoming**, 1–20.

Ryan TJ (2014) Reconsidering moral issues in politics. *The Journal of Politics* **76**(2), 380–97.

Sagi E and Dehghani M (2014) Measuring moral rhetoric in text. *Social Science Computer Review* **32**(2), 132–44.

Shepsle KA (1972) The strategy of ambiguity: Uncertainty and electoral competition. *American Political Science Review* **66**(2), 555–68.

Skitka LJ and Bauman CW (2008) Moral conviction and political engagement. *Political Psychology* **29**(1), 29–54.

Smith KB et al. (2017) Intuitive ethics and political orientations: Testing moral foundations as a theory of political ideology. *American Journal of Political Science* **61**(2), 424–37.

Sterling J and Jost JT (2018) Moral discourse in the Twitterverse: Effects of ideology and political sophistication on language use among US citizens and Members of Congress. *Journal of Language and Politics* **17**(2), 195–221.

Turner-Zwinkels FM et al. (2021) Conservatives' moral foundations are more densely connected than liberals' moral foundations. *Personality and Social Psychology Bulletin* **47**(2), 167–184.

Van Leeuwen F and Park JH (2009) Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences* **47**(3), 169–73.

Voelkel JG and Feinberg M (2018) Morally reframed arguments can affect support for political candidates. *Social Psychological and Personality Science* **9**(8), 917–24.

Wang SYN and Inbar Y (2021) Moral language Use by US political elites. *Psychological Science* **32**(1), 14–26.