

Moral Foundations of U.S. Political News Organizations

William E. Padfield¹ & Erin M. Buchanan, Ph.D.²

¹ Missouri State University

²

Author Note

William Padfield is a master's degree candidate in psychology at Missouri State University. This thesis partially fulfills the requirements for the Master of Science degree in Psychology.

Correspondence concerning this article should be addressed to William E. Padfield, 901 S. National Ave, Springfield, MO, 65897. E-mail: Padfield94@live.missouristate.edu

Abstract

The media ecosystem has grown, and political opinions have diverged such that there are competing conceptions of objective truth. Commentators often point to political biases in news coverage as a catalyst for this political divide. The Moral Foundations Dictionary (MFD) facilitates identification of ideological leanings in text through frequency of the occurrence of certain words. Through web scraping, the researchers extracted articles from popular news sources' websites, calculated MFD word frequencies, and identified words' respective valences. This process attempts to uncover news outlets' positive or negative endorsements of certain moral dimensions concomitant with a particular ideology. In Experiment 1, the researchers gathered political articles from four sources. They were unable to reveal significant differences in moral or political endorsements, but they solidified the method to be employed in further research. In Experiment 2, the researchers will expand their number of sources to 10 and will analyze articles that pertain to the 2018 confirmation hearings of U.S. Supreme Court Justice Brett Kavanaugh. This topic was selected due to the moral disputes associated with his nomination.

Keywords: politics, morality, psycholinguistics

27

Moral Foundations of U.S. Political News Organizations

28 In the United States, today's media landscape affords consumers a multitude of options
29 for obtaining political news. Since the advent of cable news networks and the World Wide
30 Web in the last decades of the twentieth century, consumers have gained access to an
31 ever-expanding menagerie of news sources, many of which can be called up via a simple click,
32 touch, or swipe. Concurrent with this growth in available news sources, concerns regarding
33 political bias in news reporting have entered public consciousness. For example,
34 commentators argue that networks including Fox News Channel and MSNBC communicate
35 political news from a conservative and liberal slant, respectively. These purported biases
36 have been a cause for concern given the potential for incomplete or inaccurate news
37 reporting potentially resulting from these biases. Given the inherently moral nature of many
38 political arguments and positions, bias in news reporting might manifest as differing moral
39 appeals. Specifically, the use of differing moral language in political articles might be an
40 indicator of political bias in news media.

41 Morality and ethics have been of interest to thinkers, academics, and philosophers since
42 antiquity. Starting chiefly in the twentieth century, a scientific approach to humans'
43 understanding of morality emerged under the domain of psychology. Theories attempting to
44 explain the development and application of people's moral intuitions built the foundation for
45 the subfield of moral psychology. As the field developed, however, considerable debate has
46 taken place regarding operational definitions of "morality." Concerns regarding
47 operationalization remain an issue in the field in the twenty-first century as researchers
48 attempt to infer moral and political leanings from text and speech.

49 Moral Foundations Theory

50 As a discipline, modern moral psychology started in the late 1960s with Lawrence
51 Kohlberg (Haidt & Graham, 2007). Kohlberg's research popularized his theory of the
52 development of moral reasoning. This theory establishes the steps of moral reasoning through
53 which humans proceed as their cognitive structures assume higher levels of sophistication and
54 nuance (Kohlberg & Hersh, 1977). Kohlberg borrowed from Jean Piaget's stages of cognitive
55 development in which children progress from the sensorimotor through to the formal
56 operations stage. Similarly, Kohlberg found people typically start with a "pre-conventional"
57 understanding of morality during infancy in which children understand "right" and "wrong"
58 purely in terms of how they interact with resultant experiences of rewards and punishment.
59 Typically, people progress through several steps until they reach a "post-conventional" ethics.
60 People who have reached the post-conventional stage are said to be able to weigh competing
61 abstractions and reason their way to a conclusion that promotes justice based upon their
62 "self-chosen ethical principles" (Kohlberg & Hersh, 1977). From Kohlberg's perspective,
63 issues of justice and fairness comprise the foundation of morality (Haidt & Graham, 2007).
64 This view persisted until it encountered criticism in the early 1980s.

65 Kohlberg's conception of morality faced major scrutiny from psychologist Carol
66 Gilligan. In 1982, Gilligan criticized Kohlberg's theory on the grounds that it focused solely
67 on the moral concerns of men, and that it ignored those of women (Haidt & Graham, 2007).
68 Gilligan drew attention to purported differences in the ways men and women are taught to
69 relate to self and others. She offered a historic argument contending women have
70 traditionally filled roles related to caring and nurturing. She pushes back against Kohlberg's
71 assumption that moral development replaces "rule of brute force," as enforced by men, with
72 the justice-based "rule of law." According to Gilligan, this assumption implies women are
73 less morally developed, owing to their absence both in masculine displays of violence as well
74 as in enforcement of the law (Gilligan, 1982). Gilligan argues for the existence of a distinct,

but equal development process that women and girls must undergo in order to develop their moral selves. Stark differences in the ways women are traditionally taught to interact with their social world cause them to develop ethical systems based upon their non-aggressive relationships with others. Gilligan thus asserted morality was built upon an alternative moral foundation: caring (Gilligan, 1982). This debate between competing conceptions of morality did not resolve until Gilligan and Kohlberg conceded the existence of two moral foundations: justice and caring (Haidt & Graham, 2007). While this new direction in moral psychology appeared to represent a more inclusive outlook on the construct, these novel ideas would soon be challenged on the grounds of its apparent western-centric outlook.

Jonathan Haidt and Jesse Graham formulated Moral Foundations Theory as a method by which to capture the entirety of humans' moral domain (Haidt & Graham, 2007). The researchers argued older theories of moral psychology were focused primarily on issues of justice, fairness, and caring - individually focused foundations of morality that align with the beliefs of political liberals (Haidt & Graham, 2007). In other words, moral psychology ignored the valid moral foundations of conservatives. Moral Foundations Theory (MFT) holds that people's moral domain can be mapped by quantifying their endorsement of five moral foundations: *harm/care*, *fairness/reciprocity*, *ingroup/loyalty*, *authority/respect*, and *purity/sanctity* (Haidt & Graham, 2007).

In their brief overview of the history of moral psychology, Graham, Haidt, and Nosek (2009) explained Shweder, Much, Mahapatra, and Park's objections to moral psychology as it stood in the late 1980s. Their criticism centered on the fact moral psychology concerned itself with issues regarding justice and individuals' rights. Such a system, they argued, did not account for moral concerns outside of the western world (Graham et al., 2009). Individually focused concerns can be grouped under an overarching "ethic of autonomy," which was thought to be one of three ethics upon which humans base moral decisions. The other two ethics were the "ethic of community" (comprising one's duty to their family, tribe,

etc.), and the “ethic of divinity” - representing one’s duty not to defile their God-given body and soul (Graham et al., 2009). In the 2000s, Haidt and Graham (2007) took this line of reasoning further in their assertion that moral psychology favored certain political ideologies over others.

Haidt and Graham settled on these specific foundations after the completion of a literature survey of research in anthropology and evolutionary psychology (Graham et al., 2011). The researchers attempted to locate virtues and morals corresponding to “evolutionary thinking.” For instance, the researchers cited Mauss’ work on reciprocal gift-giving, which informed the establishment of the *fairness/reciprocity* foundation. Additionally, evolutionary literature on disgust and its correlation to human behavior regarding food and sex informed the *purity/sanctity* foundation (Graham et al., 2011). The researchers identified the five “top candidates” for the foundations of human cultures’ morality (Graham et al., 2011).

The first two foundations (*harm/care* and *fairness/reciprocity*) are termed the “individualizing foundations,” as they are centered on the concerns of individuals rather than groups. *Harm/care* represents an endorsement of compassion and kindness, while opposing cruelty and harm. *Fairness/reciprocity* represents concerns centered on guaranteeing individual rights as well as justice and equality among all people. The other three foundations (*ingroup/loyalty*, *authority/respect*, and *purity/sanctity*) are the “binding” foundations, owing to their focus on group-related concerns, rather than those of individuals. *Ingroup/loyalty* represents endorsements of patriotism and heroism and discourages nonconformity and dissent. *Authority/respect* represents an endorsement of social hierarchies and traditions while denigrating disobedience. Finally, *purity/sanctity* represents concerns regarding chastity and piety, while discouraging vices and indulgences, including lust, avarice, and gluttony (Haidt & Graham, 2007). Liberals tend to endorse the individualizing foundations more than conservatives. Conservatives, on the other hand, tend to endorse the

binding foundations more than liberals. It should be noted, however, conservatives also tend to endorse all five foundations equally, implying they base moral judgments on all foundations (Graham et al., 2009).

Moral Foundations Dictionary

In order to capture language's role in moral and political reasoning, Graham et al. (2009) formulated the Moral Foundations Dictionary (MFD) in order to capture moral reasoning and justification as used in speech and text. The MFD is composed of 259 words, with around 50 words assigned to each of the five foundations. The researchers created a preliminary list of words that they believed would be associated with the five foundations. Then, using the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, & Frances, 2007) computer program, they analyzed transcripts of liberal and conservative Christian sermons in order to obtain frequencies of the occurrence of words from the researchers' initial list. The researchers manually checked the results from LIWC in order to make sure the results make sense given the contexts and rhetorical devices used in the sermons, as word frequency analysis ignores sentence context. The researchers offered the following example from a Unitarian sermon as a demonstration of ambiguous statements requiring human verification: "Don't let some self-interested ecclesiastical or government authority tell you what to believe, but read the Bible with your own eyes and open your heart directly to Jesus" (Graham et al., 2009). This sentence added to the *authority/respect* total in LIWC's analysis, but it appears to suggest that one should reject authority in this context. The researchers eliminated this sentence from the *authority/respect* raw count on account of this discrepancy between the use of authority-related words and the speaker's clear intentions (Graham et al., 2009).

Similar to previous research on Moral Foundations Theory, liberal ministers used *harm*, *fairness*, and *ingroup* words more often than conservative ministers. Conversely, conservative

ministers used *authority* and *purity* words more often than liberal ministers. However, conservative ministers did not use *ingroup/loyalty* words more than liberals. Rather, liberal ministers used words pertaining to *ingroup/loyalty*, but in contexts that promote rebellion and independence - causes *opposite* to positive endorsements of that foundation (Graham et al., 2009).

To this point, most text analysis utilizing the Moral Foundations Dictionary operationalizes endorsement of any one of the foundations as percent occurrence of words in a given text from the foundation's respective word list. As such, most analyses assume that zero percent occurrence is indicative of no endorsement, while any non-zero percent occurrence indicates endorsement of the foundation. This operational definition may not be sufficient in describing the true nature of the writer or speaker's endorsement of one of the sets of moral intuitions. A quick glance at the MFD words for *harm/care* reveals the presence of words that are more closely associated with universally accepted conceptions of *harm* over *care* and vice-versa (Graham et al., 2009). For example, the word "cruel" has relatively negative connotations compared to "benefit." For the *harm/care* foundation, it is conceivable that use of the word "cruel" might indicate a greater attentional focus of the idea of *harm* rather than *care*.

For *harm/care*, the definition of the foundation, as well as its name, clearly distinguishes between two somewhat opposite sides of an attentional continuum, with *harm* on the negative end and *care* on the positive side. In other words, the entries in the MFD for *harm/care* have somewhat clear positive and negative valences. The same pattern can be seen in the MFD entries for the other four foundations. *Purity/sanctity* features words that likely have a negative valence to most observers, including "disease" and "trash," along with more positive words, including "right" and "sacred" (Graham et al., 2009). These dichotomies, however, bring up other questions regarding the definition and names of the other four foundations apart from *harm/care*: *fairness/reciprocity*, *ingroup/loyalty*,

authority/respect, and *purity/sanctity*. The latter four foundations have names that are harder to understand as a valence continuum, as the concepts in the names are more similar, even to the point of being virtually synonymous in the case of *fairness/reciprocity*.

When considering the issue of positive versus negative valence in MFD words, the question of how texts are analyzed vis-a-vis the MFD remains. How can raw percentage of MFD word occurrence capture the valence and focus of the writer or speaker? If 2% of a politician's speech features positive words (i.e., "benefit" and "defend") from the MFD *harm/care* list, how can researchers be sure the level and nature of the speaker's "endorsement" of the foundation equals that of another politician whose speech contained negatively connoted MFD words from the *harm/care* list? They would have equal endorsements as far as the numbers are concerned, but the words used and focus given are on opposite sides of the *harm/care* spectrum.

This issue is compounded by the fact the Moral Foundations Questionnaire (MFQ) and its subscales assume endorsement lies on a continuum. The Moral Foundations Questionnaire (MFQ), which was developed subsequent to the MFD, measures individuals' endorsements of each of the foundations using a six-point scale (Graham et al., 2011). The questionnaire is made up of judgment items and relevance items. Judgment items are phrased such that the respondent signals their agreement or disagreement with straightforward statements. An example of such a statement reads: "It can never be right to kill another human being" (Graham et al., 2011). Relevance items gauge the respondent's opinion regarding the importance of foundation-related concerns. For example, the respondent is directed to rate how important the following situation is to their sense of morals: "whether or not someone did something disgusting." This example measures the relevance of the *purity/sanctity* foundations. Each foundation has a judgment and relevance subscale, totaling 10 subscales for the MFQ (Graham et al., 2011).

The aforementioned ambiguity of the Moral Foundations Dictionary as an instrument

becomes clearer upon closer examination of the items in the Moral Foundations Questionnaire. One item under the *fairness/reciprocity* judgment subscale reads, “Justice is the most important requirement for a society” (Graham et al., 2011). The survey respondent must select a number on a scale from 1 to 6 indicating responses spanning “strongly disagree” at 1 to “strongly agree” at 6. While the scales in the MFQ do not represent true valence as it pertains to individual words, it does allow for a greater degree of specificity in terms of an individual’s endorsement of a particular moral foundation. When a respondent selects a 4 for the aforementioned MFQ statement, they clearly are indicating they “slightly agree” with the statement (Graham et al., 2011). This specificity is not present in most analyses involving the MFD and percent occurrence, unless they also take into account the valence of the words used in the text or speech of interest.

Valence

Borrowing from Osgood’s work in the 1950s, Bradley and Lang (1999) recognized valence as one of three related dimensions comprising emotion when developing their Affective Norms for English Words (ANEW). As mentioned before, “valence,” the first dimension, denotes the pleasantness of a given word. “Arousal,” the second dimension, describes the stimulating nature of a word. Lastly, “dominance” or “control” describes the extent to which a word makes one feel in or out of control (Bradley & Lang, 1999). The researchers developed ANEW by presenting participants with a list of 100-150 words and asking for them to rate the word on all three dimensions using the Self-Assessment Mannikin (SAM), which allows ratings along either a nine-point scale when using traditional paper instruments or a twenty-point scale when using a computerized version.

Participants saw the stimulus word and responded on each scale. The valence scale featured a smiling figure at one end (representing pleasantness) and a frowning figure at the other end (for unpleasantness). The arousal scale had a “wide-eyed” figure at one end with a

sleepy figure at the other, representing stimulating and unstimulating respectively. Finally, the dominance scale featured a large figure, indicating the highest degree of control, at one end and a small figure, indicating a lack of control, at the other end (Bradley & Lang, 1999). The end result of this procedure yielded affective norms along the three dimensions for 1,040 English words (Bradley & Lang, 1999). ANEW represented an important first step in establishing affective norms for large numbers of English words. However, later researchers found the 1,040-word list to be limiting for a language consisting of thousands of words.

Warriner, Kuperman, and Brysbaert (2013) exponentially lengthened the list of words with affective norms to 13,915 English lemmas, the base forms of words without inflection (i.e., “watch” rather than “watched” and “watching”). The researchers recognized the importance of affective norms in several areas of study, including emotion, language processing, and memory (Warriner et al., 2013). They argue the list of words included in ANEW is sufficient for small-scale factorial research designs, but the list is “prohibitively small” for larger-scale “megastudies” that are common in psycholinguistic research today (Warriner et al., 2013).

In order to source a large number of lemmas for affective ratings, the researchers drew from several validated sources. These include the 30,000 lemmas with age-of-acquisition (average age at which a particular word is learned) ratings gathered by Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012) as well as the content lemmas from the SUBTLEX-US corpus consisting of subtitles from various forms of visual media (New, Brysbaert, Veronis, & Pallier, 2007). This data collection resulted in the final list of 13,915 lemmas. Lists of 346-350 words were presented to participants recruited through the Amazon Mechanical Turk subject pool. Participants rated the words along one of the three dimensions, unlike the ANEW project in which participants rated each word along all three dimensions at once. The researchers used a nine-point scale similar to the one used by Bradley and Lang (1999) when collecting ratings for ANEW (Warriner et al., 2013).

The researchers noted several points of interest upon observing ratings. First, they found that valence and dominance ratings had a negative skew, indicating more words elicited feelings of happiness and control than their respective opposites. Also, when examining the relationship between valence and arousal ratings, the researchers found a U-shaped relationship. This U-shape indicates words with high degrees of positivity and negativity elicited higher arousal (Warriner et al., 2013). These observations along with the now-greatly expanded list of affective norms has been applied to several lines of inquiry in psycholinguistics.

Warriner and Kuperman (2015) utilized the new affective norms list in order to investigate the validity of the Pollyanna hypothesis, or the prevalence of a generally optimistic outlook in humans as reflected in language. The researchers were able to conclude the existence of a greater number of positive-valence English words in the list of 13,915 lemmas. Additionally, after observing token frequency in a number of text corpora, including SUBTLEX-US, the Corpus for Contemporary American English (COCA), the British National Corpus (BNC), Touchstone Applied Science Associates, Inc. Corpus (TASA), and the corpus used for the Hyperspace Analogue to Language model (HAL), the researchers found that words with positive valence were also used more frequently (Warriner & Kuperman, 2015). While the researchers concede the possibility of an acquiescence bias in ratings as a possible explanation for the observed positivity bias, this investigation represents one application of the Warriner et al. (2013) list in emotional studies.

In addition to applications in emotion research, the Warriner et al. (2013) norms have been utilized in cognitive research as well. One cognition-based study investigates the relationship between emotion and response latencies in word recognition. Kuperman, Estes, Brysbaert, and Warriner (2014) sought to use these new norms to fill in the knowledge gaps regarding variance in word recognition. The researchers drew several conclusions regarding emotion and word recognition (specifically in naming and lexical decision tasks - two

cognitive processing tasks wherein a participant has to read aloud or judge a word for its
lexicality). First, Kuperman et al. (2014) found slower decision-making and reading times in
negative-valence words, faster times in neutral words, and even faster times in words with
positive valence. The researchers also concluded that words causing higher arousal tend to
have slower decision times than less-arousing words. They found valence had a stronger
effect on recognition than arousal (both effects were independent, not interactive). They
found an interaction between emotion and word frequency such that valence and arousal are
more effective on lower frequency words than high frequency words. Finally, Kuperman et al.
(2014) found a greater effect of valence and arousal on response latency for lexical decision
tasks than for naming tasks (Kuperman et al., 2014). This research serves as further
evidence that the Warriner et al. (2013) list can be used for research inquiries both within
and without the field of psycholinguistics.

In the present studies, the researchers used the Warriner et al. (2013) list in order to
denote the valence of the words appearing in the news articles scraped from the internet.
Valence was considered as another independent variable and its relationship with the words
comprising the Moral Foundations Dictionary were of chief interest to the researchers. The
valence was used as a means to determine whether individual words in the MFD represented
more positive aspects of their respective foundation or if they denoted a more negative
aspect of the foundation. Specifically, valences were used to weight the MFD words by their
relative degree of positivity or negativity. Incorporating word valence into a study involving
the MFD is meant to alleviate some of the issues regarding the aforementioned ambiguity
regarding the words in the Moral Foundations Dictionary.

News Media and Politics

Research into politics, language, and media has illuminated the complex relationships
between all three. Any politically-oriented discussion of word occurrence as an implication of

moral or political position assumes that language and ideology are intrinsically linked. Deborah Cameron (2006) points out the expressive nature of ideological beliefs and how that expression is conveyed through language, thus implying a connection between ideology and language. She goes on to criticize the notion that language is either the “pre-existing raw material” used to shape ideologies or the “post-hoc vehicle” for their propagation. Rather, the structure of language itself is shaped by ideology and social processes even when it is used to explain or express ideologies (Cameron, 2006). Owing to the fact the Moral Foundations Dictionary was developed in order to assess the moral, which includes the ideological, orientation of discourse, its purported ability to assess parts of the structure of language (vocabulary) for ideological lean is of chief interest to the researchers in the present study.

The use of language both to express and further an ideological goal has been documented in the techniques employed by candidates for political office in the U.S., Druckman, Jacobs, and Ostermeier (2004) considered political “issues” as communication that attempts to persuade constituents to vote for the candidates based on their strengths in matters of public policy. According to the researchers, “image” priming describes techniques deployed in order to sway votes based on favorable aspects of the candidate’s behavior and personality (Druckman et al., 2004). The researchers investigated political issue and image priming on the part of candidates as implied by the disproportionate attention candidates paid to particular issues over others. The researchers found numerous examples of issue and image priming during the 1972 re-election campaign of Richard Nixon.

They linked the Nixon administration’s awareness of the issues for which the president had public support to the issues he should emphasize (and prime) during the campaign. Likewise the researchers found evidence that Nixon’s team was aware of negative evaluations of his warmth and trustworthiness, and thus took steps to prime his purportedly positive qualities, including strength and competence (Druckman et al., 2004). The researchers also cited research from Iyengar and Kinder (1987) suggesting the news media affected

perceptions of President Jimmy Carter's competence by emphasizing (e.g., priming) issues related to energy, defense, and the economy. This focus implies news media may contribute to Americans' perception of politicians based on where the media places emphasis.

There is a potential caveat regarding the validity of Druckman et al. (2004)'s findings: reproductions of several studies purporting to demonstrate social priming effects have failed to replicate the original results. Pashler, Coburn, and Harris (2012) point out the distinction between perceptual and social (or goal) priming both in their operational definitions as well as their replicability. Perceptual priming often works through the inducement of a certain response from a related prime, as in, for example, semantic priming. Social (or goal) priming encompasses phenomena by which people exhibit complex behavioral changes subsequent to exposure to a prime. Pashler et al. (2012) point out well-known studies investigating social priming, including the use of elderly-related primes to induce slower walking speeds in participants. Studies investigating perceptual priming have been "directly replicated in hundreds of labs" (Pashler et al., 2012). This replication rate does not appear to be the case for social priming, as argued by Pashler et al. (2012).

Pashler et al. (2012) noticed the unusually large effect size values (Cohen's d) reported by researchers studying social priming effects. The researchers reproduced two studies from Williams and Bargh (2008). The first study attempted to prime participants by having them plot points on a Cartesian grid. The independent variable was priming condition and contained three levels: short, middle, and long distance. Those instructed to plot points further apart were hypothesized to express a higher degree of psychological distance regarding their family. The second study used the same priming conditions, but hypothesized that greater distance between points would prime participants to estimate fewer calories in unhealthy foods than those who were primed with shorter distances between points. Pashler et al. (2012) concluded those two studies from Williams and Bargh (2008) held little validity while also casting doubt on the prevalence of social priming effects themselves, based on the

inability of other researchers to replicate previously reported effects in this area.

While these concerns regarding the replication of social priming studies are valid and deserve further investigation, Druckman et al. (2004) does not purport to demonstrate a widespread effect of social priming on the American electorate. In other words, this research makes no claim to empirically supported priming effects. Rather, Druckman et al. (2004) chronicle the efforts on the part of the Nixon Administration to prop up the president's supposed strengths while downplaying his weaknesses. These tactics were deployed through the careful use of language in order to achieve the administration's political goals. As such, Druckman et al. (2004)'s research on Nixon serves as an example of language's potential utility in the propagation of desirable political opinions. The researcher's investigation of news media's focus on specific issues during the Carter Administration likewise provide an example of language as a potential conduit for the transfer of politically biased information. The idea that even 1970s news media could contain political biases is of particular interest to the current study, which investigates similar phenomena in contemporary news media.

Other research into news media suggests certain media outlets, at least indirectly, may have an effect on the voting records of representatives in Congress (Clinton & Enamorado, 2014). Specifically, the researchers identified a pattern of declining support for President Bill Clinton's policies chiefly among Republicans in the House of Representatives after the Fox News Channel began broadcasting on cable and satellite systems in their respective districts. As Fox News was, at the time of its launch in 1996, the only outwardly ideological national news network, the researchers were able to track its spread across the country and observe voting records of members of Congress both before and after Fox News' arrival. The researchers concluded that members of Congress, excluding those newly elected at the time of Fox News Channel's emergence, attempted to anticipate resultant conservative-leaning shifts among their constituents by bolstering their conservative voting record before the next election (Clinton & Enamorado, 2014).

Therefore, the current study sought to combine both methods related questions and extension/replication of previous moral foundation results found for liberal and conservative sources. First, the MFD was combined with previous research by the current authors (see below) and weighted by valence to create weighted percentages to better specify endorsement. Second, these weighted percentages were examined for their differences in across liberal and conservative news sources.

Experiment 1

Method

For Experiment 1, the researchers approached the study with the intention to answer a method question. That is, this portion of the current research was conducted in order to solidify the best method by which to analyze political news text under the Moral Foundations Theory framework while also alleviating some of the aforementioned valence problem observed in the Moral Foundations Dictionary. The researchers hypothesized the news sources generally perceived as liberal leaning (*NPR* and *The New York Times*) would contain MFD words and valences indicating endorsements of the individualizing moral foundations (*harm/care* and *fairness/reciprocity*). Additionally, the researchers hypothesized the two sources generally perceived to be conservative leaning (*Fox News* and *Breitbart*) would feature MFD words and valences indicating equal endorsement of all five foundations.

Sources

Political articles were collected from the websites of four notable U.S. news sources, a process known as web scraping. The sources were *The New York Times*, *National Public Radio (NPR)*, *Fox News*, and *Breitbart*. They were selected for their widespread recognition and the fact they are easily categorized (by the general public) according to perceived

political lean. In general, *The New York Times* and *NPR* are perceived by many to have a liberal bias or lean. In contrast, Fox News and Breitbart are believed to have a conservative bias or lean (Mitchell, Matsa, Gottfried, & Kiley, 2014). Political articles in particular were identified and subsequently scraped by including the specific URL directing to each source's political content in the *R* script. For example, rather than scrape from nytimes.com, which would return undesired results (non-political features, reviews, etc.), we instead included nytimes.com/section/politics so that more or less exclusively political content was obtained. All code for this manuscript can be found at <https://osf.io/5kpj7/>, and the scripts are provided inline with this manuscript written with the *papaja* library (Aust & Barth, 2017).

Identification of the sources' political URLs presented a problem for two of the sources owing to complications with how their particular sites were structured. While in the multi-week process of scraping articles, we noticed word counts for *NPR* and *Fox News* were not growing at a similar pace as those from *The New York Times* and *Breitbart*. Upon investigation, we found another, more robust URL for political content from NPR: their politics content "archive." The page structure on NPR's website was such that only a limited selection of articles is displayed to the user at a given time. Scraping both the archive and the normal politics page ensured we were obtaining most (if not all) new articles as they were published. We later ran a process in order to exclude any duplicate articles. *Fox News* presented a similar issue. We discovered *Fox News* utilized six URLs in addition to the regular politics page. These URLs led to pages containing content pertaining the U.S. Executive Branch, Senate, House of Representatives, Judicial Branch, foreign policy, and elections. Once again, duplicates were subsequently eliminated from any analyses.

Materials

Using the *rvest* library in the statistical package *R*, we pulled body text for individual articles from each of the aforementioned sources (identified using CSS language) and

compiled them into a dataset (Wickham, 2016). Using this dataset, we identified word count and average word count per source. This process was run once daily starting in February 2018 until March 2018. Starting in mid-March 2018, the process was run twice daily - once in the morning and again in the evening. Data collection was terminated once 250,000 words per source was collected in April 2018.

Data analysis

Once data collection ended, the text was scanned using the *ngram* package in *R* (Schmidt, Gonzalez-Cabrera, & Tomasello, 2017). This package includes a word count function, which was used to remove articles that came through as blank text, as well as to eliminate text picked up from the Disqus commenting system used by certain websites. At this point, duplicate articles were discarded.

The article text was processed using the *tm* and *ngram* packages in *R* in order to render the text in lowercase, remove punctuation, and fix spacing issues (Feinerer & Hornik, 2017). The individual words were then reduced to their stems (i.e., *abused* was stemmed to *abus*). The same procedure was applied to the MFD words and the words in the Warriner et al. (2013) dataset. Using the Warriner et al. (2013) dictionary, the words making up each of the five foundations in the MFD were matched to their respective valence value.

Concurrent research by Jordan, Buchanan, and Padfield (2019) is assessing the validity of both the Moral Foundations Questionnaire and the Moral Foundations Dictionary through a multi-trait multi-method analysis of the two instruments using multiple samples. The instruments and foundation areas are being analyzed against one another, in order to test reliability, as well as against the Congressional Record in order to test predictive validity for political orientation. The researchers were able to identify a number of potential new words that, if added to the MFD, could comprise a dictionary with greater validity, and less

likelihood of zero percent texts, as this often occurs with the current MFD. Those results have informed this analysis, and their updated findings may change the underlying dictionary used in this analysis (albeit, we do not expect any changes in the results presented below).

The source article words were compiled into a dataset where they were matched up with their counterparts in the MFD along with their valence and a percentage of their occurrence. Therefore, for each article, the percentage of the number of *harm/care* words occurring in the articles were calculated, and this process was repeated for each of the foundations. Words' percent occurrence were multiplied by their *z*-scored valence. Valences were *z*-scored in order to eliminate any ambiguity regarding the direction of the valence. Positive values indicate positive valence, and negative values indicate negative valence. Words were categorized in accordance to their MFD affiliation, creating a weighted sum for each moral foundation.

Results

To analyze if news sources adhered to differences in word use based on their target audience, we utilized a multilevel model (MLM) to analyze the data. MLM is a regression technique that allows one to control for the repeated measurement and nested structured of the data, which creates correlated error (Gelman, 2006). Using the *nlme* library in *R* (Pinheiro, Bates, Debroy, Sarkar, & Team, 2017), each foundation's weighted percentage was predicated here by the political lean of the news source, using the individual news sources as a random intercept to control for the structure of the data.

The multilevel model did not indicate the presence of any significant or practical effect of political lean for any of the five moral foundations. The strongest effect size was observed for the *authority/respect* foundation, but the effect was in the opposite direction from what was originally hypothesized - liberal sources tended to use more *authority/respect* words than

did conservative sources. Descriptive and test statistics, p -values and effect sizes (Cohen's d) can be found in Table 1. To interpret the weighted scores, one can examine the mean and standard deviations for each. A zero score for the mean, with a non-zero standard deviation, would indicate a perfect balance of positive and negative words in each category, likely representing a neutral tone when all words are considered. Negative percentages would indicate more representation of the negative words in the MFD area, while positive percentages indicate an endorsement of the positive words in a MFD. Therefore, we suggest using the sign of the mean score to determine the directionality of the endorsement for the MFD (positive, neutral, negative), and the standard deviation to ensure that a zero score is not zero endorsement (i.e., a SD of zero indicates no words were used). Based on the weighted percent values for the five foundations, the researchers observed that MFD words seem to make up a small portion of the article text. Furthermore, the observed percentages and means appear to indicate a generally positive endorsement of all five foundations across both liberal and conservative sources.

Discussion

The results obtained in Experiment 1 did not confirm the hypothesis. The researchers found little compelling evidence of an effect of partisan lean on MFD endorsement. The strongest effect found was for the *authority/respect* foundation owing to the fact its Cohen's d value was greater than the other four foundations. However, the effect was in the opposite direction of that which was hypothesized. Specifically, the results indicated that liberal leaning sources demonstrated higher positivity regarding that foundation than conservatives. This result is contrary not only to the research hypothesis for Experiment 1 but also to previous findings in Moral Foundations Theory research. It should be noted, however, the effect size was small and the relationship was not found to be statistically significant.

Upon speculation, the researchers identified one possible reason for why the results

were unable to confirm the hypothesis. The selection of the broad and amorphous topic of “political news” may have led to the scraping of large numbers of articles with little to no moral-centric content. Rather, many articles may have been, for example, simple reporting on congressional procedures that would leave little room for the use of moral language here, let alone words from the Moral Foundations Dictionary. In short, the range of topics covered in Experiment 1 was likely too broad. The possibility exists that a tighter focus on one political issue or event, especially one that (on the surface) has a stronger relationship with morality might be more illuminating for research in moral language in news media.

Owing to the exploratory nature of Experiment 1, the researchers were afforded the opportunity to consider changes to the method to be utilized in Experiment 2. Generally speaking, the researchers believe their methodology to be sound. Web scraping methods and text processing remain viable methods for collecting large amounts of text and subsequently rendering that text in a form suitable for data analysis. Experiment 1 also demonstrated a method by which to address inherent problems in the Moral Foundations Dictionary relating to valence. The solution provided in Experiment 1 appears to provide insights into the MFD words where none previously existed. Finally, calculating weighted percentages and sums for each moral foundation provides an easily interpreted summary of MFD word positivity and occurrence.

While the methodology used in Experiment 1 features many strengths, there are aspects which could be strengthened for future studies. The researchers identified two such changes that were subsequently employed in Experiment 2. First, the researchers elected to include more news sources for web scraping and analysis in addition to the four used in Experiment 1. Second, the researchers chose to focus their data collection efforts exclusively on one event in U.S. politics: the nomination and confirmation of Justice Brett Kavanaugh to the U.S. Supreme Court. In Experiment 2, the researchers sought to confirm the usefulness and validity of the method as well as test a similar hypothesis as Experiment 1.

Experiment 2

Kavanaugh Supreme Court Hearing

In the wake of Justice Anthony Kennedy's retirement from the Supreme Court of the United States, President Donald Trump nominated Brett Kavanaugh as the new Associate Justice. Kavanaugh was previously on the U.S. Court of Appeals for the District of Columbia. The Senate Judiciary Committee began his confirmation hearing on September 4, 2018 (US Government, 2018a). Following allegations of sexual assault by high school classmate Dr. Christine Blasey Ford, the committee postponed its vote on whether or not to open the confirmation to the entire Senate.

On September 27, the committee questioned Dr. Ford before commencing a second round of questioning for Judge Kavanaugh (US Government, 2018b). During the intervening weeks between hearings, two more women came forward with two separate allegations of sexual assault on the part of Kavanaugh. According to Nielsen reports, more than 20 million people watched the September 27 proceedings on television (O'Connell, 2018). This figure does not take into account viewers who watched online, nor does it account for viewers outside the United States. On September 28, the Senate Judiciary Committee voted to send the nomination to the Senate floor. Senator Jeff Flake of Arizona, however, lobbied for a week-long FBI investigation on Kavanaugh and the allegations facing him, which the committee, and later the President, approved. The investigation concluded with no significant findings. The Senate voted 50-48 to approve Kavanaugh's appointment on October 6, 2018 (US Government, 2018c).

The Kavanaugh nomination, confirmation hearing, and eventual swearing-in, as well as the news media's coverage of all three events, feature many moral dimensions that likely differ depending on one's morals. On one side of the debate, Kavanaugh's Supreme Court tenure presents a prime opportunity to bring morality back into interpretation of the

Constitution. Kavanaugh’s confirmation creates a conservative stronghold among the justices on the court. Commentators have noted this might help advance a judicial agenda that backpedals certain rights previously upheld by the Supreme Court, including abortion and gay marriage - social issues challenged by their opponents at least partially on moral grounds. On the other side of the debate, the assault allegations have energized Kavanaugh’s opponents to advocate for his rejection from the court owing to misdeeds resulting from Kavanaugh’s own alleged lack of morals. Additionally, the moral duty of the Senate as the upper chamber in the U.S. legislature has been scrutinized in public discourse with respect to its handling of the assault allegations vis-a-vis Kavanaugh’s confirmation.

In contrast to Experiment 1, the researchers approached Experiment 2 with the intention to confirm the method employed was valid for the analysis of the scraped text as well as for any inferences drawn from the analyses. For Experiment 2, the researchers hypothesized that news sources perceived as liberal will exhibit positive endorsements of the individualizing moral foundations (*harm/care* and *fairness/reciprocity*) in their articles reporting on the Kavanaugh confirmation hearing. News sources perceived as conservative are hypothesized to positively endorse all five foundations equally in their coverage of the Kavanaugh hearing. The researchers tested the hypothesis by analyzing the content scraped from news sources’ web pages during the two weeks before and two weeks after Kavanaugh’s confirmation, owing to its prominence in the news. The content will be analyzed for valence and moral alignment under Moral Foundations Theory.

Method

Sources

Articles pertaining to the Brett Kavanaugh Supreme Court nomination and confirmation were scraped from the websites of 12 U.S. news sources. These sources were

selected owing to their favorability among political partisans according to Mitchell et al. (2014). The sources favored by liberals were *The New York Times*, *National Public Radio* (NPR), *Slate*, *Huffington Post*, and *Politico* (Mitchell et al., 2014). The sources favored by conservatives included *Fox News*, *Breitbart*, *The Rush Limbaugh Show*, *The Blaze*, and *Sean Hannity*. Political articles referencing Brett Kavanaugh’s nomination process were identified and subsequently scraped by including the URL for each source’s coverage of the nomination in the *R* script. All code for this manuscript can be found at <https://osf.io/5kjp7/>, and the scripts, again written with the *papaja* library in *R*, are provided inline with this manuscript (Aust & Barth, 2017).

Materials

Using the *rvest* library in the statistical package *R*, we pulled body text for individual articles from each of the aforementioned 10 news sources (identified using CSS language). We compiled the articles into a dataset (Wickham, 2016). Using this dataset, we identified word count and average word count per source. This process was run for articles pertaining to Kavanaugh’s nomination that were published between September 13, 2018 and October 11, 2018 inclusive. This date range was selected in reference to the widely-publicized and viewed nomination hearing on September 27, 2018. We set the start date at September 13 (two weeks before the hearing) and the end date at October 11 (two weeks after the hearing) so that we could capture a large amount of data (roughly one month) during which Kavanaugh’s nomination was at its peak saturation in news coverage.

Data analysis

As in Experiment 1, the text was scanned with *ngram*. Again, blank articles, text from the Disqus system, and duplicate articles were removed (Schmidt et al., 2017). The text was

processed and stemmed in order to convert to a usable form for further analysis (Feinerer & Hornik, 2017). Words were subsequently matched with their valences from Warriner et al. (2013). Depending on the results of Jordan et al. (2019)’s research, alternative forms of the Moral Foundations Dictionary may be imported instead of the original dictionary.

Using the *tm* and *ngram* packages in *R*, the researchers processed the text in order to convert it to lowercase, fix spacing anomalies, and remove punctuation (Feinerer & Hornik, 2017). Each individual word was reduced to its stem (i.e., *diseased* was stemmed to *diseas*). Once again, the same procedure was applied to the MFD words and the words in the Warriner et al. (2013) dataset. Using the Warriner et al. (2013) dictionary, the words in the MFD were assigned their respective valence. The researchers obtained the words’ percent occurrence in the text. Once again, percents were multiplied by *z*-scored valence and categorized into their proper MFD category.

Experiment 2

Results

Discussion

Conclusions

References

Aust, F., & Barth, M. (2017). papaja: Create APA manuscripts with R Markdown.

Retrieved from <https://github.com/crsh/papaja>

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW):*

Instruction Manual and Affective Ratings (No. C-1). The Center for Research in

Psychophysiology, University of Florida.

Cameron, D. (2006). Ideology and language. *Journal of Political Ideologies*, 11(2), 141–152.

doi:10.1080/13569310600687916

Clinton, J. D., & Enamorado, T. (2014). The National News Media's Effect on Congress:

How Fox News Affected Elites in Congress. *The Journal of Politics*, 76(4), 928–943.

doi:10.1017/S0022381614000425

Druckman, J. N., Jacobs, L. R., & Ostermeier, E. (2004). Candidate Strategies to Prime

Issues and Image. *The Journal of Politics*, 66(4), 1180–1202.

doi:10.1111/j.0022-3816.2004.00295.x

Feinerer, I., & Hornik, K. (2017). Text mining package. Retrieved from

<http://tm.r-forge.r-project.org/>

Gelman, A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do.

Technometrics, 48(3), 432–435. doi:10.1198/004017005000000661

Gilligan, C. (1982). New maps of development: New visions of maturity. *American Journal*

of Orthopsychiatry, 52(2), 199–212. doi:10.1111/j.1939-0025.1982.tb02682.x

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different

sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5),

1029–1046. doi:10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2), 366–385. doi:10.1037/a0021847

Haidt, J., & Graham, J. (2007). When Morality Opposes Justice: Conservatives Have Moral Intuitions that Liberals may not Recognize. *Social Justice Research*, 20(1), 98–116. doi:10.1007/s11211-007-0034-z

Jordan, K. N., Buchanan, E. M., & Padfield, W. E. (2019). *A Validation of the Moral Foundations Questionnaire and Dictionary*. Retrieved from <https://osf.io/kt9yf/>

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory into Practice*, 16(2), 53–59. doi:10.1080/00405847709542675

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, 143(3), 1065–1081. doi:10.1037/a0035669

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44(4), 978–990. doi:10.3758/s13428-012-0210-4

Mitchell, A., Matsa, K. E., Gottfried, J., & Kiley, J. (2014). Political Polarization & Media Habits | Pew Research Center. Retrieved from <http://www.journalism.org/2014/10/21/political-polarization-media-habits/>

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28(4), 661–677.

doi:10.1017/S014271640707035X

O’Connell, M. (2018). Ford-Kavanaugh Ratings: Hearing Brings 20 Million Viewers to Cable and Broadcast | Hollywood Reporter. Retrieved from <https://www.hollywoodreporter.com/live-feed/ford-kavanaugh-ratings-hearing-brings-20-million-viewers-cable-broadcast-1147785>

Pashler, H., Coburn, N., & Harris, C. R. (2012). Priming of Social Distance? Failure to Replicate Effects on Social and Food Judgments. *PLoS ONE*, 7(8). doi:10.1371/journal.pone.0042510

Pennebaker, J. W., Booth, R. J., & Frances, M. E. (2007). *Liwc2007: Linguistic inquiry and word count*. Austin, TX.

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., & Team, R. C. (2017). nlme: Linear and nonlinear mixed effects models. Retrieved from <https://cran.r-project.org/package=nlme>

Schmidt, M. F., Gonzalez-Cabrera, I., & Tomasello, M. (2017). Children’s developing metaethical judgments. *Journal of Experimental Child Psychology*, 164, 163–177. doi:10.1016/j.jecp.2017.07.008

US Government. (2018a). Congressional Record. *Congressional Record*, 164(146), 46. doi:10.1097/00017285-197507000-00018

US Government. (2018b). Congressional Record. *Congressional Record*, 164(160), 93. doi:10.1097/00017285-197507000-00018

US Government. (2018c). Congressional Record. *Congressional Record*, 164(167), 14. doi:10.1097/00017285-197507000-00018

Warriner, A. B., & Kuperman, V. (2015). Affective biases in English are bi-dimensional.

Cognition and Emotion, 29(7), 1147–1167. doi:10.1080/02699931.2014.968098

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191–1207. doi:10.3758/s13428-012-0314-x

Wickham, H. (2016). Package ‘ rvest ’.

Williams, L. E., & Bargh, J. a. (2008). Keeping One ’ s Distance. *Psychological Science*, 19(3), 302–308. doi:10.1111/j.1467-9280.2008.02084.x

Table 1

Experiment 1 Results - Multilevel Model

Foundation	M_C	SD_C	M_L	SD_L	t	p	d
Harm/Care	0.50	2.21	0.49	2.21	-0.21	.850	0.01
Fairness/Reciprocity	1.13	1.38	1.11	1.38	-0.42	.715	0.02
Ingroup/Loyalty	1.28	1.63	1.34	1.63	0.30	.789	-0.04
Authority/Respect	0.72	1.62	1.06	1.62	3.17	.087	-0.20
Purity/Sanctity	1.11	1.48	1.27	1.48	2.37	.141	-0.09

Note. For mean and standard deviation values, 'C' and 'L' refer to 'conservative' and 'liberal,' respectively