Measuring Moral Rhetoric in Text

Eyal Sagi[1] and Morteza Dehghani[2]

[1]Kellogg School of Management, Northwestern University

[2]Brain and Creativity Institute, University of Southern California

Abstract

In this paper we present a computational text analysis technique for measuring the moral loading of concepts as they are used in a corpus. This method is especially useful for the study of online corpora as it allows for the rapid analysis of moral rhetoric in texts such as blogs and tweets as events unfold. We use Latent Semantic Analysis to compute the semantic similarity between concepts and moral keywords taken from the Moral Foundation Dictionary. This measure of semantic similarity represents the loading of these concepts on the five moral dimensions identified by Moral Foundation Theory. We demonstrate the efficacy of this method using three different concepts and corpora.

Keywords: moral rhetoric, Moral Foundations theory, discourse analysis, Latent Semantic Analysis, Moral Foundations dictionary

Introduction

The goal of the proposed computational text analysis technique is to track and measure transformations in moral concerns with regards to different social-cultural issues, and give insights into the moral dimensions of different debates and conflicts. We show that our method can be used to track and analyze moral rhetoric and explore dynamics of various moral intuitions. Moral rhetoric can be defined as the language used for advocating or taking a moral stance towards an issue by invoking or making salient various moral concerns. The use of this type of rhetoric is prevalent in political debates and discourse (Marietta, 2002; Clifford & Jerit, in press) and it has been argued that moral rhetoric affects people's moral and political worldviews and increases levels of

political intensity (Lakoff, 2008; Marietta, 2008; Dehghani, Iliev, et al. 2009; Koleva, Graham, Iyer, Ditto & Haidt, 2012).

In addition to giving insights about various political and social conflicts, our computational method for analyzing moral rhetoric provides an alternative approach for studying morality in a more natural setting compared to self-report survey methods and artificial paradigms used in traditional judgment and decision-making experiments. Given the vital role of social networking and blogging in various political/cultural/consumer campaigns, such tools can provide us the ability to track responses to these events as they naturally unfold, allowing longitudinal analysis of changes in different moral intuitions and other psychological factors. The present project extends contemporary attempts for detecting linguistic features related to emotions and moral concerns, by coupling a computational text analysis approach for detection of frames with a prevalent theory of morality in psychology. Specifically, we ground our computational approach in the Moral Foundations Theory (Haidt & Joseph, 2004; Graham et al., 2013). The Moral Foundations Theory posits that there are five psychological systems or intuitions that account for various aspects of morality. This theory argues that each of these five foundations serve different but related social functions and the degree of sensitivity towards these foundations vary across different cultures and context. Moreover, these foundations are innate but the sensitivity towards them can change over time. Specifically, the Moral Foundations Theory has focused on the following five kinds of moral concerns (each covering both virtues and vices):

1. Care/harm: Prompted by concerns about caring and protecting individuals from harm.

2. Fairness/cheating: Concerns triggered by acts of cooperation, reciprocity and cheating.

3. Loyalty/betrayal: Related to virtues of patriotism, self-sacrifice and loyalty, and the vice of betrayal, unfaithfulness and disloyalty to the group.

4. Authority/subversion: Prompted by concerns about obedience, respect, insubordination or subversion for authority.

5. Purity/degradation: Related to the emotion of disgust and triggered by practices related to sanctity, degradation and pollution.

The Moral Foundations Theory has been used to investigate and better understand various political attitudes in the US. Particularly, Graham, Haidt and Nosek (2009) demonstrate that liberals and conservatives attend to different moral intuitions: while liberals focus exclusively on the notions of harm and fairness when making moral judgments, conservatives also attend to ideas of loyalty to in-group members, authority and purity. More recently, Koleva, Graham, Iyer, Ditto & Haidt (2012) demonstrated that the Moral Foundations Theory can explain various dimensions of the "culture war" in the US. Specifically, they show that endorsement of five moral foundations predicted support for various political issues (e.g., abortion, immigration, same-sex marriage, death penalty, gun control) more accurately than factors such as ideology, religiousness and gender, among others.

Measuring Moral Dimensions

Similar to the majority of other approaches in moral psychology, the degree of endorsement of the above intuitions described in the Moral Foundations Theory are generally measured using a self-reported questionnaire titled The Moral Foundations

Questionnaire (Haidt, Graham & Joseph, 2009). This questionnaire is now widely used and has become a standard scale in the field of moral psychology (Graham et al., 2013). Subsequently, Graham et al. (2009) produced The Moral Foundations Dictionary (MFD) to be used for text analysis purposes. This dictionary consists of 295 words and word stems related to each of the moral intuitions of harm, fairness, authority, loyalty to in-group and purity. The MFD has been used as part of the Linguistics Inquiry and Word Count (LIWC) software (Pennebaker, Francis & Booth, 2003; Tausczik & Pennebaker, 2010) to investigate differences in moral concerns between different cultural groups (e.g. Graham, Haidt & Nosek, 2009). For example, Graham, Haidt and Nosek (2009), use the MFD and LIWC to analyze sermons delivered in conservative and liberal churches. Consistent with their previous experiments in which they used the Moral Foundations Questionnaire, they show that sermons given at liberal churches tend to focus more on issues related to harm and fairness compared to sermons given at conservative churches. In contrast, conservative churches tend to focus more on issues related to authority and purity. Clifford and Jerit (in press) use the MFD for performing a manual text analysis of twelve years of coverage in the *New York Times* focusing on the debate over stem cell research. They show that advocates and opponents of stem cell research focus on different moral concerns, and the use of moral rhetoric increases during periods of legislative activity. However, given that their analysis was done manually, it was limited only to the domains of harm and purity. Dehghani, Sagae, Sachdeva and Gratch (2013) examine the differences between liberal and conservative moral value systems using a hierarchical generative topic modeling technique based on Latent Dirichlet Allocation (LDA; Blei, Ng, & Jordan, 2003) to enable the unsupervised detection of topics in their

corpus of liberal and conservative weblogs. They use small sets of words selected from the MFD as seeds to encourage the emergence of topics related to different moral concerns, and examined similarities and differences in how such concerns are expressed between these groups. Consistent with findings in moral psychology, they demonstrate that there are significant differences in how liberals and conservatives construct their moral belief systems.

Discourse Analysis

Many of the methods for measuring moral dimensions, such as those described above, use an entire text as their unit of measurement. That is, they provide a measure of moral loading for each text or group of texts. However, texts are rarely limited to a single topic. Moreover, even in texts that are concerned with a single central theme, other themes are often introduced. Likewise, discussions of a specific topic often involve its relation to other topics. For example, when discussing the war on drugs, it is common to also discuss money laundering and other types of criminal activities, as well as addiction and its consequences to society. Consequently, in this paper we describe a method that measures the moral dimensions for a particular topic, rather than a text as a whole. In order to achieve this level of sub-document focus we take our cue from research in discourse analysis and focus on measuring the moral loadings associated with particular concepts through the keywords that are associated with them.

In discourse analysis, researchers often focus on discovering specific keywords that identify the important points of the discourse. One measure used for this purpose is keyness, the frequency of a term in the document in relation to its expected frequency (cf.

Scott & Tribble, 2006). For example, Baker (2004) uses keyness to discover words of importance in the debates over the gay male law reform in the U.K. House of Lords.

Another pattern of word occurrence that is often employed in discourse analysis is that of collocation. At its core, two words are collocated if they appear together in the text (often as immediate neighbors). The more likely two words are to collocate the more related their meaning is thought to be (Firth, 1957). For example, Baker (2004) uses the collocates of his target terms of interest, gay and homosexual to determine that, in the debates he is exploring, the term gay is often associated with identity while the term homosexual is often associated with specific sexual acts.

Over the last few decades, computational models and statistical methods have been developed that harness the power of these types of manual analysis to provide some insights into the semantic relatedness of words. Perhaps the best known of these models is Latent Semantic Analysis (Deerwester et al., 1990; Dumais & Landauer, 1997; Schütze, 1998) which uses the patterns of word co-occurrence (i.e. collocations within the same document) to generate a semantic space which can be used to measure the semantic similarity of words. The method we describe below uses such a semantic space as the foundation upon which measure of moral rhetoric is built.

General Method

While keyword-based methods like LIWC can be used to measure the moral rhetoric over entire documents, we are interested in measuring the rhetoric related to specific topics. The method we describe in this paper focuses on the words that are used in relation to the topic or topics of interest. To do this we adapt the approach described by Sagi, Diermeier and Kaufmann (in press). In particular we measure the similarity of the

contexts in which specific keywords appear to the contexts in which words from the

Moral Foundation Dictionary appear. Our measure of moral loading is then based on this

similarity - The more similar the contexts of a particular keyword to those of the words

associated with a specific moral dimension, the higher the loading on that dimension.

Figure 1 provides a schematic overview of the method, which is described in more detail

below. In essence, we replace the simple frequency count approach used by LIWC with a

measure of semantic similarity.

Please insert Figure 1 about here.

At its most basic level, this method is an application of Latent Semantic Analysis,

an approach that quantifies the meaning of words as vectors in a high dimensionality

space (Deerwester et al., 1990; Dumais & Landauers, 1997; Schütze, 1997). This space is

generated by tabulating the co-occurrence pattern of words – The particular method we

used, Wordspace (Schütze, 1997, 1998; Takayama, et al., 1998; Infomap, 2007) uses a

word-by-word matrix in which each row is associated with a keyword and each cell in

that row represents the frequency with which a particular word appears within a pre-

determined window before or after the keyword (we used ±15 words). A generalized

factoring method, Singular Value Decomposition (SVD) is then used to transform this

matrix into a semantic space in which only the N most important dimensions are used (in

this paper, N=100). The closer two vectors are within this space, the more similar their

meaning is thought to be. The most common measure for this distance is cosine similarity – the cosine of the angle between the vectors. Following the recommendation of Hu, et al. (2007), we discard the first dimension prior to computing similarity because this dimension is always positive and correlates with the underlying terms' frequency in the corpus. If both vectors are of equal magnitude (as is the case when the vectors are normalized), this measure is essentially the correlation between the vectors.

However, we are interested not in the overall similarity between word meanings in a corpus, but in their similarity in parts of the corpus – such as during a particular period, or for a particular speaker or speakers. For this purpose, we make use of another property of the semantic space – Namely, that it is possible to calculate the vector representing the meaning of a group of words, such as a sentence or a paragraph, by using vector addition over the content-bearing words that appear within that group. The resulting context vector is an approximate representation of the topic or gist of that group of words. Consequently, we can judge how a word is used in a portion of a corpus by looking at the context vectors representing the contexts in which it appears in that portion of the corpus.

After we have computed the context vectors for all appearances of our topical keywords and the morally-relevant keywords in the portion of the corpus that we are interested in, we can measure the similarity of these two groups of contexts. We measure this similarity by randomly sampling pairs of vectors representing contexts – one from each group and without repetition and calculating their cosine similarity. This is repeated until the pool of available contexts in one of the groups is exhausted. Next, we average all of the calculated cosines and generate an overall measure of similarity between the two

groups. To ensure the stability of the result, we repeat this process 1,000 times and use the average of the resulting similarity measures. The end product of this process is a single number representing the similarity of the set of contexts in which a topical keyword appears to the set of contexts in which a morally-relevant keyword appears. The more similar these sets of contexts are, the more likely it is that the specific morally-relevant word is relevant to the topic, in a fashion similar to the frequency counts achieved by LIWC.

However, our method differs from LIWC on several key aspects. Most importantly, whereas LIWC measures the frequency of terms within a document, we measure the semantic similarity between the terms identifying a moral dimension and the terms identifying our concept of interest. Consequently, our measure does not rely on *how often* a moral term appears, but rather on how similar *its use* is to that of the key term. In many cases, moral terms that are important for the topic will also appear more frequently in tandem with it. However, that is not necessarily the case. In that sense, our measure is a generalization and an improvement over LIWC.

Case studies

We use our text analysis technique to analyze the rhetoric in three different socio-political conflicts to gain better insights into the moral dynamics of them. Specifically, we will focus our attention on the following issues: the moral dynamics surrounding the World Trade Center before and after the '93 and '01 attacks, the conflict over the "Ground-Zero Mosque", and the abortion debate in the US Senate. In all these cases we demonstrate that our method can be used to examine and gain insight into the moral dimensions of the conflict. Moreover, by relying on three different types of corpora

(blogs, articles, speeches) we show that the described method is a general purpose method that can be used on various different types of text.

World Trade Center vs. Empire State Building

Introduction

For our first case study we chose to explore the effect of the two terrorist attacks on the World Trade Center (WTC below) in '93 and '01 had on the manner in which it was described. We predict that each attack resulted in a dramatic rise in the moral rhetoric surrounding the building. Moreover, since terrorist attacks are usually associated with the harm dimension, we expect that the change in moral rhetoric would focus on that dimension. However, the change in moral loading following these attacks should be confined to the building itself and not affect discussions of other topics. To test this we picked a control term that is also a landmark building in New York City – The Empire State Building (ESB below).

Corpus description (NYT)

The corpus used in this analysis is a collection of all the articles printed in the New York Times from January, 1987 to June, 2007 (Sandhaus, 2008). It includes about 1.8 million articles totaling approximately 5.8 billion words.

In this and the following analyses, we used all of the moral keywords from the MFD that occur in this corpus over 10,000 times (231 words overall).

Results

Figure 2 shows the overall moral loading over time for the two terms. As predicted, paired-samples t-tests revealed that there was a sharp increase in the overall moral loading of the World Trade Center for both 1993 ($t(4) = 4.61$, $p < .01$) and 2001

($t(4) = 6.61$, $p < .01$). No such effect is seen for the moral loading of the Empire State Building (1993: $t(4) = 2.46$, $p = 0.07$; 2001: $t(4) = 0.24$, $p = 0.82$). In fact, the moral loading of the ESB exhibited a slight decrease from 1992 to 1993. The figure also demonstrates another interesting result – Following the dramatic increase in moral rhetoric in '93 and '01 the level of rhetoric gradually decreased over time (Pearson's correlations: $r_{93-'00} = -.64$, $p = .09$; $r_{01-'07} = -.997$, $p < .0001$). This suggests that the terrorist attacks made a lasting, though not permanent, change in the moral rhetoric journalists used when writing about the WTC.

Please insert Figure 2 about here.

Perhaps more interesting is the change in the relationship among the dimensions following the attacks (Figure 3). To analyze these changes we computed the mean moral loading for each dimension by year. We then used a regression model in which the three periods and the five different moral dimensions were independent variables. Year was entered as a random variable. Cosine distance was used as the dependent variable. As expected, there was an increase in moral loading of the WTC over time ($F(2, 89) = 195.67$, $p < .0001$). The moral loading of the ESB did not differ significantly over time ($F(2, 89) = 0.47$, $p = 0.63$). Both the WTC and the ESB exhibited an overall difference in loading among the 5 dimensions (ESB: $F(4, 89) = 24.88$, WTC: $F(4, 89) = 81.92$, both ps $< .0001$). Most importantly, the profile for the WTC exhibited a significant change in the

relative importance of the dimensions over time as evidenced by a significant interaction

between the time and dimension variables ($F(8, 89) = 13.34$, $p < .0001$). The most salient

dimensions affecting this interaction were *harm* and *loyalty* to in-group (both ps $< .01$).

In contrast, no such interaction was found for the ESB ($F(8, 89) = 0.26$, $p = 0.98$). The

pattern observed is consistent with our hypothesis that the harm moral rhetoric would

increase in use following a terrorist attack. Overall, these results suggest that major

events can precipitate lasting change in moral rhetoric.

Please insert Figure 3 about here.

Ground Zero Mosque

Introduction

The conflict surrounding the Cordoba Muslim Community Center in NYC has

been one of the most controversial political issues in US politics in the last several years.

This issue became morally significant, for both conservatives and liberals in the US, in a

short period of time when it started getting framed as the "Ground-Zero Mosque". During

the conflict, there was a prevalent use of moral rhetoric by both groups, clashing the

sacred American value of religious freedom against the moral decadence of

contamination of the "hallowed ground" (Davis & Dover, 2010) at Ground Zero. This

debate was started on the blogosphere by a single blogger (Elliott, 2010), and most of the

discussions regarding this issue took place on various different political blogs. This

provided us a representative sample of the rhetoric used in the conflict and also the ability to track responses to events as they naturally unfolded (Dehghani, Sagae, Sachdeva, Gratch, 2013).

Corpus description (Ground Zero Mosque Blogs)

We compiled blog posts (and comments) related to the Ground-Zero Mosque from popular 10 popular conservative and liberal weblogs that were posted between January 1, 2010 and December 31, 2010 (for details about this corpus please see Dehghani, et al., in press). This corpus contains a total of 3449 blog posts, consisting of 1575 posts from the conservative blogs and 1874 from liberal blogs.

Results

Our main hypothesis in this analysis was that moral loadings with respect to the ground zero mosque would be higher during the debate than either before it started or after it ended. Consequently, we predicted that an inverted-U shaped function would be a good predictor of the moral loadings over the time course of the debate.

To test this prediction we computed the average cosine distances between each MFD keyword and the term "mosque" using a semantic space based on the New York Times. We then computed the moral loading of "mosque" by averaging these mean distances for the liberal and conservative blogs for each month. Figure 4 plots the resulting moral loading measure by month.

Please insert Figure 4 about here.

Using a regression analysis we compared a linear model in which month was the dependent variable and a U-shaped model in which the square of the month was the dependent variable. The U-shaped model had its center set in mid-June (the middle of the dataset). Because both models had a single free parameter, they could be compared directly. The U-shaped model outperformed the linear one for both liberals (U-shaped: r^2 = 0.48, *RMSE* = .019, *p* = .0074; Linear: r^2 = 0.045, *RMSE* = .025, *p* = .25) and conservatives (U-shaped: r^2 = 0.41, *RMSE* = .018, *p* = .015; Linear: r^2 = 0.22, *RMSE* = .020, *p* = .072).

It is also possible to take a more in-depth look and examine how the moral loading of "mosque" was changed by the debate (Figure 5). We conducted an analysis of variance where the moral dimension and debate period (January-February, March-September, and October-December) were the independent variables and the measured moral loading was the dependent variable. The analysis indicated that the debate involved all 5 dimensions of the MFD equally. There was an overall increase in loading ($F(2, 45)$ = 75.85, *MSE* = 0.0001, *p* < .0001), but no interaction between the individual dimensions and the increase over time ($F(8, 45)$ =1.59, *MSE* = 0.0001, *p* = .15). Likewise, individual analyses indicated that all 5 dimensions show an increase over time (all *p*s < .01).

Please insert Figure 5 about here.

These results suggest that the measure of moral loading we present in this paper can be used as a general measure tracking the level of moral loading used by participants in a debate.

The abortion debate in the US Senate

Introduction

The topic of our third, and final, analysis regards the moral positions taken by democratic and republican senators on the issue of abortion. Abortion is an issue that divides U.S. society along party lines – On one side of the debate, Democrats argue that the most important consideration when it comes to abortion is the woman's right to make choices that affect her life. On the other side, Republicans hold that the fetus's right to life is paramount. Most frequently, these two positions are represented in the media as pro-choice and pro-life, respectively. Sagi, Diermeier and Kaufmann (2013) demonstrated that these positions are reflected in speeches given in the U.S. Senate. However, this leaves the question of the moral underpinnings of these positions open. In this analysis we attempt to delve a little deeper and explore the types of moral rhetoric the different parties bring to bear on the debate. In particular, we are interested in the moral aspects of abortion that guide each party's position.

Corpus description

For this analysis we used transcripts of U.S. Senate speeches collected by Bei Yu (Yu et al., 2008). This corpus includes all speeches given on the Senate floor and spans the years 1989-2006. It totals approximately 180 million words spanning almost 230,000 speeches.

Results

As before we computed the average cosine distances between each MFD keyword and our target keyword. In this study we used the term 'abortion'. Interestingly, Republicans used this term more frequently than Democrats (n_D = 3807; n_R = 7147). This contrasts with terms such as pregnancy (n_D = 1972; n_R = 1231) and pregnant (n_D = 1670; n_R = 1052), which were used more often by Democrats. Consequently, there was a significant interaction in the frequency of use of the terms (Chi-square test: $X^2(2)$=1112.97, $p < .0001$). This suggests that Republicans discuss the topic of abortion more frequently than Democrats do. One possible reason that Republicans raise this topic more frequently than Democrats, especially in a political forum such as the U.S. Senate, is that this topic is of particular moral importance to them. Because abortion is such a polarizing topic in the U.S., we might expect that the perceived importance of the topic will correlate with its moral loading. Consequently, we hypothesized that Republicans would show a higher moral loading on this topic.

As before, this analysis was based on the average distances between abortion and each moral dimension by year. As expected, Republicans exhibited a higher overall moral loading than Democrats (paired-samples t-test: M_D = 0.017, SD_D = 0.004; M_R = 0.022, SD_R = 0.006; $t(4) = -3.81$, $p = .019$). In fact, Republicans exhibited a higher moral loading than Democrats on all 5 MFD dimensions.

The most highly loaded dimensions overall were fairness and purity, and therefore we focused on these dimensions for further analysis. In particular, our measure of moral loading demonstrated that, in the case of abortion, Democrats were most concerned with issues of fairness whereas Republicans were most concerned with issues that relate to purity (Figure 6; Analysis of Variance: $F(1, 17) = 6.59$, $MSE = 0.26$, $p = .020$).

Please insert Figure 6 about here.

These differences in rhetoric seem reasonable as they align with public (and media) perceptions of each party's position with respect to abortion. Democrats tend to frame the debate in terms of choice (cf. Sagi, et al, in press), a term that is closely associated with a discourse of rights and fairness. Surprisingly, the dimension of harm, associated with the common republican frame of life, is only the third-highest loaded dimension for Republicans. However, a closer inspection of the specific keywords that comprise the dimension of purity suggests that the reason Republicans load so highly on it is that it contains terms such as abstinence, celibacy, and prostitution. These results suggest that while Republicans openly endorse the value of life when debating abortion, they are more frequently concerned with the relationship between abortion and sexual purity.

Discussion

There has been an upsurge in sentiment analysis research in computer science (e.g. Liu, 2012), specifically in sentiment research in social networks (e.g. Agarwal et al., 2011). The methods used, however, are for the most part not grounded in psychological theories, and rely merely on word co-occurrences. On the other hand, theories of morality in psychology have been almost entirely developed and tested using self-report surveys, and given that different theories rely on different sets of questionnaire, it is very difficulty

to judge between them and to device a unifying theory of moral cognition. The method described in this paper goes beyond both self-reported measurements and contemporary sentiment analysis techniques by coupling a natural language processing technique with a distinct theoretical emphasis on the underlying psychological factors that influence moral cognition, as well as how these factors unfold over time. Our approach makes it possible to identify and track social media content that reveals specific traits in a specific population, among the vast and seemingly insurmountable volume of user-generated content published daily.

More generally, there is evidence that relying on moral rhetoric results in a shift in the discussion domain from the mundane to the sacred, promoting absolutist reasoning and causing less willingness to acknowledge trade- offs (Marietta, 2008). Recent research shows that sacred values that play important roles in many intractable cultural conflicts (Atran & Ginges, 2012) can emerge from the use of sacred rhetoric (e.g. Dehghani et al., 2010). The method described in this paper allows detection and examination of dynamics of moral rhetoric, and can potentially give us insights about the moral significance of an issue and help us predict when "rational actors" morph into "devoted actors" (Atran, 2006).

Authors

Eyal Sagi (eyal@u.northwestern.edu)

Kellogg School of Management, Northwestern University

2001 Sheridan Rd, Evanston, IL 60208

Eyal is a postdoctoral fellow in the Kellogg School of Management at Northwestern University. His research interests include text comprehension and discourse analysis. He is particularly interested in the relationship between words and their textual context and how cultural differences are reflected in language use and comprehension.

Morteza Dehghani (mdehghan@usc.edu)

Brain and Creativity Institute, University of Southern California

3620A McClintock Avenue; Los Angeles, CA 90089, USA

Morteza is a research assistant professor of computer science, psychology, and the Brain and Creativity Institute (BCI) at University of Southern California. His research interests include computational social sciences and psychoinformatics. Specifically, he is interested in using AI techniques to investigate cross-cultural differences in decision making evident in various cultural products.

References

Atran, S., & Ginges, J. (2012). Religious and sacred imperatives in human conflict. *Science*, 336(6083), 855-857.

Baker, P. (2004). 'Unnatural Acts': Discourses of homosexuality within the House of Lords debates on gay male law reform. *Journal of Sociolinguistics*, *8*(1), 88-106.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *The journal of Machine Learning research*, 3, 993-1022.

Clifford, S., & Jerit, J. (in press). How word do the work of politics: Moral Foundations Theory and the debate over stem-cell research. *Journal of Politics*.

Cohen, T., Schvaneveldt, R., & Widdows, D. (2010). Reflective Random Indexing and

    indirect inference: A scalable method for discovery of implicit connections.

    Journal of Biomedical Informatics, 43(2), 240-256.

Davis, L. & Dover, E. (Aug 22, 2010). Ground Zero Mosque opponents, supporters turn

    out to demonstrate. *ABC News*. Available at:

    http://abcnews.go.com/US/Politics/ground-mosque-opponents-supporters-

    turn-demonstrate/story?id=11455698

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990).

    Indexing by latent semantic analysis. *Journal of the American society for*

    information science, *41*(6), 391-407.

Dehghani, M., Atran, S., Iliev, R., Sachdeva, S., Medin, D. & Ginges, J. (2010). Sacred

    values and conflict over Iran's nuclear program. *Judgment and Decision Making*,

    5, 7, pp. 540-546.

Dehghani, M., Iliev, R., Sachdeva, S., Atran, S., Ginges, J. & Medin, D. (2009).

    Emerging sacred values: Iran's nuclear program. *Judgment and Decision Making*,

    4, 7, pp. 930-933.

Dehghani, M., Sagae K., Sachdeva, S. & Gratch, J. (2013). Linguistic Analysis of the

    debate over the Construction of the `Ground Zero Mosque'. *Journal of*

    Information Technology & Politics.

Dumais, S. T., & Landauer, T. K. (1997). A solution to Plato's problem: The latent

    semantic analysis theory of acquisition, induction and representation of

    knowledge. *Psychological review, 104(2), 211-240.*

Elliott, J. (Aug 16, 2010). How the "ground zero mosque" fear mongering began. *Salon.com.* Available at:

http://www.salon.com/news/politics/war_room/2010/08/16/ground_zero _mosque_origins

Firth, J. (1957) *Papers in Linguistics, 1934-1951.* London, UK: Oxford University Press.

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, 1029-1046.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. H. (2013). Moral Foundations Theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55-130.

Haidt, J., Graham, J., & Joseph, C. (2009). Above and below left-right: Ideological narratives and moral foundations. *Psychological Inquiry*, 20, 110-119.

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

Hu, X., Cai, Z., Wiemer-Hastings, P., Graesser, A. C., & McNamara, D. S. (2007). Strengths, limitations, and extensions of LSA. *The handbook of latent semantic analysis*, 401-426.

Infomap [Computer Software] (2007). Stanford, CA. http://infomap-nlp.sourceforge.net/. Accessed 20 January 2013.

Koleva, S., Graham, J., Haidt, J., Iyer, R., & Ditto, P. H. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, 46, 184-194.

Lakoff, G. (2002). *Moral politics: how liberals and conservatives think*. University of Chicago Press, Chicago, IL.

Marietta, M. (2008). From My Cold, Dead Hands: Democratic consequences of sacred rhetoric. *Journal of Politics*, 70, 3:767-779.

Marietta, M. (2009). The absolutist advantage: sacred rhetoric in contemporary presidential debate. *Political Communication*, 26,4:388-411

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2003). Linguistic inquiry and word count: *LIWC2001 manual*. Mahwah, NJ: Erlbaum.

Sandhaus E. (2008). *The New York Times Annotated Corpus*. Philadelphia, PA: Linguistic Data Consortium.

Sagi E., Diermeier D. & Kaufmann, S. (2013) Identifying Issue Frames in Text. *PLoS one* 8(7): e69185. doi:10.1371/journal.pone.0069185

Schütze, H. (1997). *Ambiguity in Language Learning: Computational and Cognitive Models*. Chicago, IL: University of Chicago Press.

Schütze, H. (1998). Automatic word sense discrimination. *Computational linguistics*, *24*(1), 97-123.

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education* (Vol. 22). John Benjamins Publishing Company.

Takayama, Y., Flournoy, R., & Kaufmann, S. (1998). *Information Mapping: Concept-Based Information Retrieval Based on Word Associations*. Stanford, CA: CSLI Publications.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.

Yu, B., Kaufmann, S., & Diermeier, D. (2008). Classifying party affiliation from political speech. *Journal of Information Technology & Politics*, *5*(1), 33-48.

Sample (reduced) context of *abortion*:
"... we might be **encouraging abortion** by covering it with a **veneer** of **compassion**"
(Dan Coats, R, 09/04/1997)

Generating a semantic space

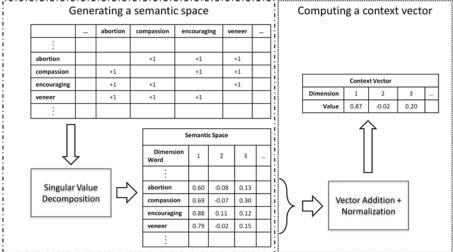|	...	abortion	compassion	encouraging	veneer	...
⋮						
abortion			+1	+1	+1	
compassion		+1		+1	+1	
encouraging		+1	+1		+1	
veneer		+1	+1	+1		
⋮						

Singular Value Decomposition ⇒

Semantic Space

Dimension Word	1	2	3	...
⋮				
abortion	0.60	-0.08	0.13	
compassion	0.69	-0.07	0.30	
encouraging	0.88	0.11	0.12	
veneer	0.79	-0.02	0.15	
⋮				

Computing a context vector

Context Vector

Dimension	1	2	3	...
Value	0.87	-0.02	0.20	

Vector Addition + Normalization

Figure 1: A schematic description of the method described in the paper. The method
involves two phases. First, a semantic space must be generated. This space is
based on a term by term co-occurrence matrix in which each column describes the
frequency with which the column term occurs within the context of the row term
(in our implementation this was a fix window of 15 words before and after the
row term). This matrix is processed using Singular Value Decomposition and the
first N dimensions of the output from the semantic space. In the second phase, the
vectors computed in the first phase are used to compute vectors representing the
context of occurrences of key terms by means of vector addition. The mean cosine
similarity between the set of vectors representing a key term and those
representing the terms describing a dimension of the MFD is then the computed
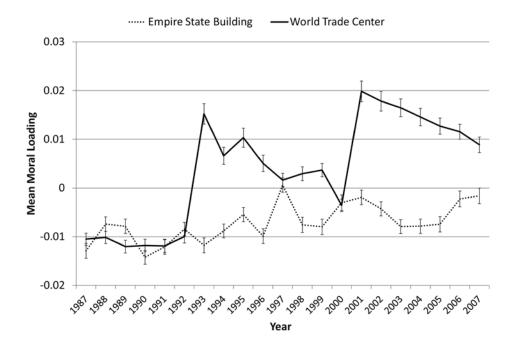moral loading of the key term on that dimension.

Figure 2: Mean moral loading for the *World Trade Center* and the *Empire State Building* over time in the NYT corpus. Error bars represent standard error.
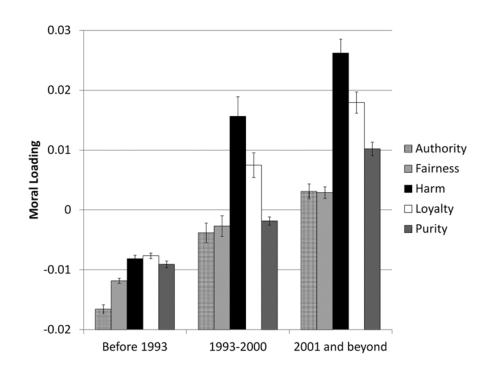
Figure 3: Mean moral loading for each MFD dimension for the *World Trade Center*

broken down by 3 periods identified by the terrorist attacks of '93 and '01.
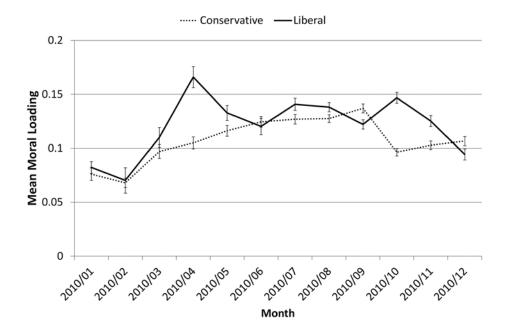
Error bars represent standard error.

Figure 4: Mean moral loading for the term 'mosque' by party affiliation over the course of the Ground Zero Mosque debate. Mean loadings were calculated on a per-month basis. The red line represents the conservative blogs and the blue line represents the liberal ones. Error bars represent standard error.
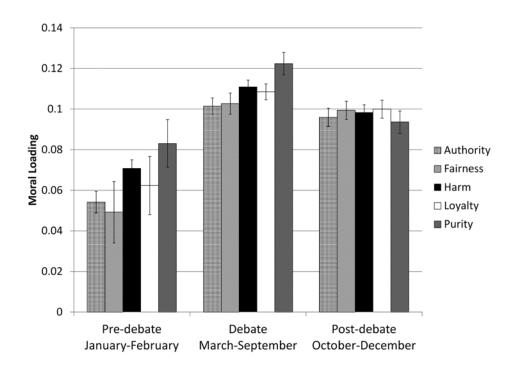
Figure 5: Mean moral loading for each MFD dimension the term 'mosque' broken down by the progress of the debate. The first period (Jan-Feb) identifies the early stages. The second period (Mar-Sep) encompasses the core of the debate. The third period (Oct-Dec) follows the debate. Error bars represent standard error.
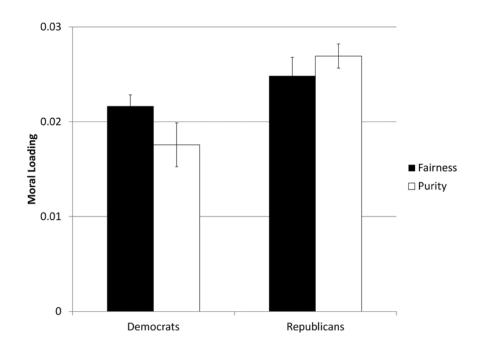
Figure 6: Moral loadings for 'abortion' by party affiliation for the MFD dimensions of *purity* and *fairness*. Error bars represent standard error.