Computational Linguistics

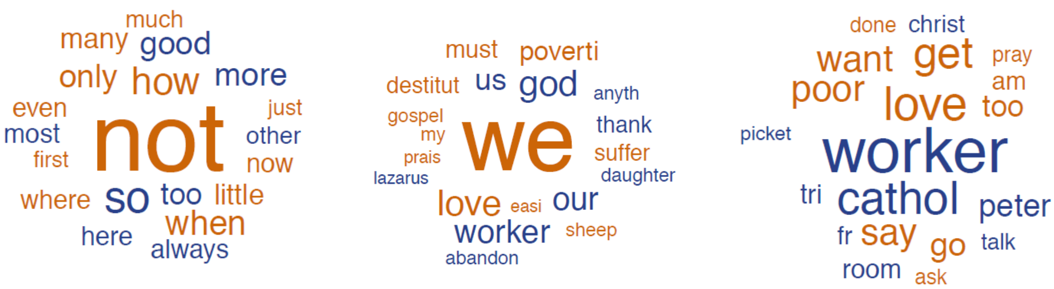
Longitudinal Analysis of Linguistic Rigidity of Value-motivated Groups

Mohammad Al Boni, Seth Green,
Megan Stiles, Katherine Harton,
Donald E. Brown
University of Virginia,
E-mail:{ma2sm, smg5b, mes5ac, klh8mr,
deb}@virginia.edu

Written text by value-motivated groups can provide insights into their way of thinking, and potentially, behaviour. Previous works by Venuti et al. and Green et al. showed that computational linguistic analysis can be performed to infer the flexibility of value-motivated groups from their writings. However, researchers have found that religious and value-motivated groups can't be analyzed collectively as they regularly evolve. To address this gap, we extend Green et al.'s method to single documents. Experimental results show that previous features used to predict groups' rigidity are less predictive for single documents' rigidity, and a newly added feature provides a significant contribution to the prediction process. Furthermore, we propose a weighting scheme for linear regression based on the inter-group variance to leverage data unbalance. Results indicate that a weighted least squares significantly outperforms a traditional least squares approach. This work supports real-time decision making, and provides a tool for practical use by practitioners.

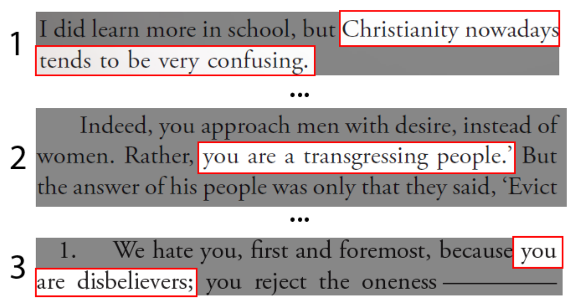
1. Introduction

The ideological divide between social, religious and political groups continue to grow, and negotiations as a way for peacemaking have been increasingly challenging. Practitioners such as intelligence analysts, politicians, and non-profit organizations require thorough understanding of ideological differences between value-motivated groups in order to develop effective negotiations strategies. Existing works performed collective linguistic analysis of written materials produced by value-motivated groups. The main premise for these works is that the use of language by groups can bring insights into their behaviour. Researchers in (Venuti et al. 2016) and (Green et al. 2017) have defined a 1-9 linguistic rigidity scale such that 1 reflects a rigid use of language while 9 reflects a flexible use. They based their inference of linguistic rigidity scores on two concepts: 1) number of contexts for value words, and 2) judgment sentences. Value words are key terms that groups use to express their beliefs and ideologies. For example, the word "hate" can be used by groups to express strong opposition towards an object. However, "hate" can be used differently by various groups as it might means immorality, wickedness, harmfulness, etc. Researchers proposed various methods for automatic extraction of value words (Venuti et al. 2016; Green et al. 2017). Figure 1 shows examples of value words extracted by previous methods. Researchers have used the number of different contexts for such value words as a measurement for the linguistic flexibility of a text. On the other hand, judgments are sentences used by groups to express an explicit opinion towards an entity (e.g., person, group, religion, etc.). Figure 2 shows examples of judgment sentences extracted from an ISIS text using methods from Venuti et al. and Green et al. Finally, the researchers in (Venuti et al. 2016) and (Green et al. 2017) have included features quantifying the sentiment of documents. The motivation for including sentiment features is to capture the extent to which positive or negative events may influence the linguistic flexibility of value motivated groups. For example, Figure 3 shows the percentage of positive words in the sermons and speeches of Martin Luther King Jr. between 1954 and 1968. We observe a general decreasing trend as a number of major events affected Dr. King and the civil right movement. We highlighted two out of the violent events in that period: the assassination of Medgar Evers on June 12, 1963, in Jackson, Mississippi and the Bloody Sunday on March 7, 1965 when state and local lawmen attacked around 600 protesters marching from Selma to Montgomery, Alabama. Also, we observed a spike in the positive sentiment and that is associated with the "I have a dream"'s speech on August 28, 1963. The main assumption for using sentiment features is that their patterns would correlate with the linguistic flexibility of groups. However, in (Green et al. 2017), Green et al. found that value words and judgment features are more effective predicting the scores of groups. In this work, we further validate this assumption for single documents.



Venuti et al. Green et al. w/pronouns Green et al. w/o pronouns

Figure 1: Examples of value words extracted by Venuti et al. and Green et al. methods.


1 I did learn more in school, but Christianity nowadays tends to be very confusing.

...

2 Indeed, you approach men with desire, instead of women. Rather, you are a transgressing people.' But the answer of his people was only that they said, 'Evict

...

3 1. We hate you, first and foremost, because you are disbelievers; you reject the oneness ——————

Figure 2: Examples of Venuti et al. (1) Green et al. (2 and 3) judgment sentences found in a snippet text from an ISIS publication.

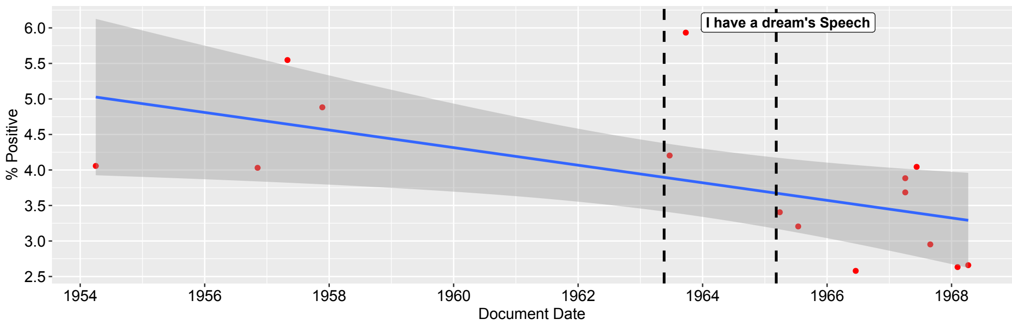


Figure 3: Percentage of words with positive sentiment in Martin Luther King Jr.'s sermons and speeches. A general decreasing trend is highlighted in blue and the two dashed lines represent the assassination of Medgar Evers and the Bloody Sunday.

While the works in (Venuti et al. 2016) and (Green et al. 2017) provide a computational tool for supporting decision making, they are at odds with evidence that value-motivated groups are constantly evolving. Therefore, a longitudinal analysis is required in order to make practical use of these method for real time decision making. In this paper, we extend Green et al.'s method to address this gap. The main contributions include: 1) applying Green et al.'s method on single documents authored by value-motivated groups; 2) expanding Green et al.'s model to include a new feature regarding the name of the group; 3) proposing a weighted least squared approach for dealing with unbalanced data. Preliminary results shows that our newly added feature has a significant contribution to the prediction process. They also show that using the inter-group variance as a weight can help boost the performance for infrequent documents at both ends of the linguistic rigidity scale. The remaining of this paper is organized as follows. Section 2 includes a literature review on theoretical frameworks and practical methods used for understanding and analyzing value-motivated groups. In section 3, we describe the data sources and pre-processing procedures used in this work. Section 4 provides the mathematical formulation of our statistical learning models along with existing and new features set. In section 5, we perform a comprehensive set of experiments to understand the predictive capability of our features and models. Finally, we conclude, in section 6, with a summary of contributions and future work.

2. Related Work

Value-motivated groups, especially religious groups, have received an extensive attention from researchers as they influence the functioning of individuals, organizations and even governments. Miller described a set of propositions on how religious organizations can influence policy making, and concluded that a longitudinal analysis is required for understanding changes in religious organizations (Miller 2002). On the other hand, Wade-Benzoni et al. have noted that values and institutions are essential for understanding ideologically-based disputes (Wade-Benzoni et al. 2002). Landrum et al. analyzed an example of ideologically-based dispute between two religious groups over climate change. Analysis of text published by both groups revealed their strategies for supporting their stances on the topic (Landrum, Tomaka, and McCarthy 2016). Other researchers such as Kröll and Strohmaier have argued that text analysis can be further used to understand human intentions (Kröll and Strohmaier 2009). Moreover, Venuti et al. have used text as a medium for analyzing ideological behavior of value-motivated groups (Venuti et al. 2016). They proposed a set of semantic and performative features and used them to estimate the linguistic rigidity of religious groups. They argued that linguistic rigidity can be used to infer the flexibility of groups which would help in policy making (i.e., starting negotiations). Green et al. extended Venuti et al.'s work to include non-religious groups (Green et al. 2017). Also, they designed new features characterizing words usages by groups (e.g., pronouns, keywords, unique words, etc.). Using the extended feature set, Green et al. were able to improve the predictive performance of linguists rigidity scores of groups.

While the works presented by Venuti et al. and Green et al. provide a tool for decision makers to analyze the flexibility of groups, they contradicts Miller's findings: religious and value-motivated groups needs to be continually analyzed to understand their strategies and organizational structure as they will change overtime. Motivated by Miller's findings, we extend Green et al.'s method to predict linguists rigidity scores of single documents.

3. Data Collection And Preprocessing

To analyze the linguistic rigidity of value-motivated groups, we require text documents written by these groups that would reflect their rigidity as it is captured by the linguistic variability. We obtained text documents from (Green et al. 2017) which included 4,568 documents (e.g., sermons, speeches, blog posts, etc.) from 16 value-motivated groups. Using the linguistic rigidity scale, researchers have manually labeled the 16 groups. They also labeled a random sample of 342 single documents. The labeling process included a number of factors such as judgment sentences and the variety of contexts in which value words are used.

We performed simple data preprocessing. Standard normalization, stopword removal (Lewis et al. 2004) and Porter stemming (Willett 2006) were applied. Next, we extracted part-of-speech tags using the nltk POS tagger (Bird, Klein, and Loper 2009). Table 1 shows the value-motivated groups sorted by their overall linguistic rigidity. The average rigidity score of single documents from most groups is close to their overall score. Also, some groups have a consistent linguistic rigidity pattern (i.e., zero or low variance across single documents). Other groups such as Unitarian and Integral Yoga have a high variance which indicates that their linguistic rigidity is changing over documents. This change can be caused by events altering the ideological behavior of groups. Figure 4 shows an example of change in linguistic rigidity pattern of the Westboro Baptist church that followed a win in the supreme court in the Snyder v. Phelps case. Prior to the ruling, documents published by the church were assigned a score of 2. Then, following the ruling in which the court found that the church's picketing of Matthew Snyder's funeral was protected under the First Amendment [1], the linguistic rigidity of the church dropped to 1. The court ruling might have embolden the church and caused its speech to become less flexible. This observation motivated this work since single document analyses can reveal insights into changes in rigidity, and allow political actors to take actions in accordance with that change (e.g., start a negotiation).

1 Facts and Case Summary - Snyder v. Phelps. Retrieved June 16, 2017, from http://www.uscourts.gov/educational-resources/educational-activities/facts-and-case-summary-snyder-v-phelps.

However, the main challenge for building computational models on single documents is the score imbalance. Figure 5 shows that documents on scores 7 and 8 constitute less than 5% of all documents which can lead to bias in the machine learning models. As discussed later in this paper, we leveraged the score imbalance issue by using the weighted least squares regression.

Group	Documents		Linguistic Rigidity Score		
	All	Labeled	Documents		Group
			Average	Variance	
ISIS	48	11	1.18	0.16	1
Pastor Anderson	228	10	1.10	0.10	1
Westboro Baptist Church	422	14	1.29	0.22	1
Malcolm X	15	12	2.08	0.08	2
Sea Shepherds	606	23	2.26	1.47	2
ACLU	40	12	3.00	0.00	3
Dorothy Day	774	18	3.72	0.80	4
John Piper	579	20	3.45	1.21	4
Rabbinic	58	58	4.00	0.00	4
Steve Shepherd	728	0	-	-	4
Bahai	73	30	4.77	1.84	6
Integral Yoga/Yogaville	59	35	5.09	2.67	6
Liberal Judaism	166	34	3.09	1.84	6
Unitarian	301	11	4.55	2.07	7
American Ethical Union	15	10	7.70	0.46	8
Meher Baba	265	44	5.52	1.98	8

Table 1: Text documents collected from 16 value-motivated groups.

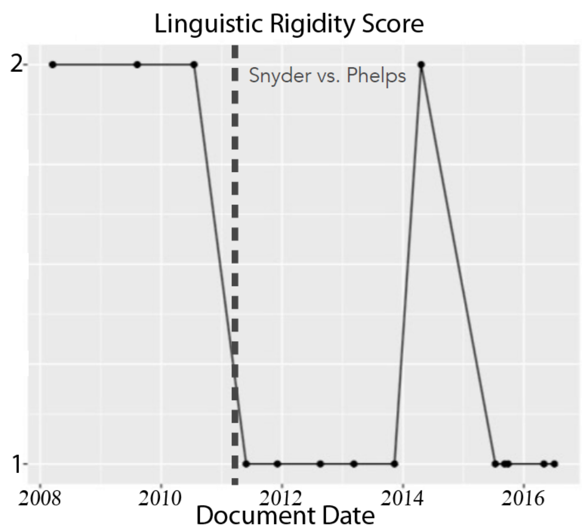


Figure 4: Linguistic rigidity scores of single documents from the Westboro Baptist Church between 2008 and 2016. A change in the rigidity is observed following the Snyder v. Phelps ruling.

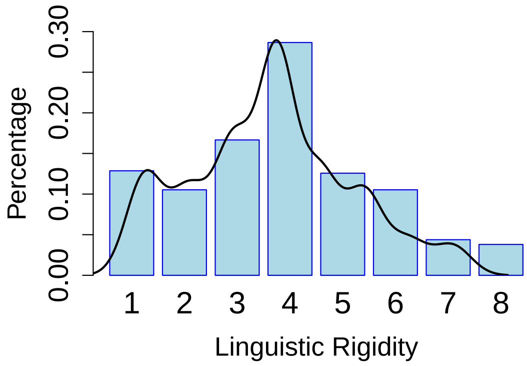


Figure 5: Percentage of documents with different linguistic rigidity scores. Fewer documents are available at both ends of the scale.

4. Mathematical Approach

We treated the problem of predicting the linguistic rigidity score of single text documents as a regression problem:

$$y_i = \beta_0 + \sum_{j=1}^{17} \beta_j * f_j(d_i) \tag{1}$$

Using this formula, we estimate the linguistic rigidity score, y_i, of document d_i using 17 features extracted from d_i. We obtained 16 features (divided into four groups) from (Venuti et al. 2016; Green et al. 2017) in addition to one newly added feature. The previous features include:

1. **Sentiment:** We obtained a lexicon of about 6,800 positive and negative words from (Hu and Liu 2004). Then, we extracted three features: 1) the percentage of positive words (perPos), 2) the percentage of negative words (perNeg), and 3) a binary variable (posDoc) about whether or not the document is positive (i.e., has more positive than negative words).
2. **Value Words:** Research has shown that analyzing the context of value words is essential for capturing the linguistic flexibility of a text (Venuti et al. 2016; Green et al. 2017). Venuti et al. proposed a method to extracting value words from text by choosing the top ten most frequent adjectives and adverbs. Green et al. argued through manual evaluation that Venuti et al. would extract only about 25% of all value words in the text. They proposed an alternative extraction method using term frequency-inverse document frequency (TF-IDF) weighting scheme (Manning, Raghavan, and Schütze 2008). For each word w in a preprocessed document (i.e., normalized, stemmed, and with stopwords removed), they count the TF of w and weight it with the IDF. Green et al. decided to calculate the IDF

using Wikipedia articles. Finally, they sorted the words with respect to their TF-IDF values and chose the top twenty words. Green et al. argued that their method were able to extract about 60% of all value words. We extracted value words using Green et al.'s method and used them to quantify two features: 1) average semantic density (avgSD) (Sagi, Kaufmann, and Clark 2009), and 2) average eigenvector centrality (avgEVC) (Estrada and Rodriguez-Velazquez 2005). The first step in calculating avgSD and avgEVC is to build context vectors of value words. A context vector is an approach for representing a word using a vector of numbers in a higher dimension that would capture the semantics of that word in different contexts. Widely used methods for creating context vectors includes word2vec (Mikolov et al. 2013), and GLOVE (Pennington, Socher, and Manning 2014). In this work, we used the GLOVE approach to represent value words co-occurrences into context vectors of length 50. Next, the avgSD is defined as the average cosine similarity between the context vectors, and it is given by:

$$avgSD(d) = \frac{1}{N}\binom{l}{2}^{-1}\sum_{i=1}^{N}\sum_{j<k}^{l}\frac{\langle c_{ij}, c_{ik}\rangle}{||c_{ik}|| \cdot ||c_{ij}||} \tag{2}$$

where N is the number of value words in document d, c_{ij} is the context vector for occurrence j of value word w_i. The purpose of using the avgSD feature is to capture the width of the semantic range of value words by averaging l of their occurrences. In another words, if value words are used in the same or similar contexts (i.e., linguistically rigid) then the pair-wise cosine similarities will have high values, and therefore, the overall avgSD will be high as well. As for the avgEVC, we used the GLOVE vectors to build a graph such that each value word is a node, and the weight of an edge linking two value words equals to the cosine similarity between the context vectors of the two value words. Next, we Search Results prune the graph by cutting any edges with weight less than 0.524. Finally, we average the eigenvector centrality (EVC) scores of all nodes. This feature would quantifies the compactness of value words since each EVC score would reflect the impact of a node by counting the number of influential connected nodes (Estrada and Rodriguez-Velazquez 2005).

3. **Judgment:** Venuti et al. have defined a judgment sentence as any sentence that starts with a noun phrase followed by "to be" verb and either an adjective or an adverb. An example of a judgment sentence, as defined by Venuti et al., is: "The book is holy". On the other hand, Green et al. proposed a different way for identifying judgments as sentences that include any value word and a pronoun. An example of a sentence that would matches Green et al. definition is: "We loved the meeting". We used both definitions and created three features: 1) the frequency of Venuti et al.'s judgment sentences (vJudgementCount), 2) the fraction of Venuti et al.'s judgment sentences (vJudgementFrac), and 3) the fraction of Green et al.'s judgment sentences (gJudgementFrac).

4. **Pronouns:** Green et al. emphasized that the use of pronouns can provide insights into the linguistic rigidity of value-motivated groups (Green et al.

2017). We used 7 features to quantify the fraction of 7 pronoun categories, as shown in Table 2, in a text document.

5. **Unique Words:** We added a feature to quantify the fraction of distinct words in a text document. This feature would reflect the breadth of vocabulary used by a value-motivated group.

6. **Group Name:** We included the name of the value-motivated group as a categorical feature. Knowing that documents are authored by the same group, this feature will help in preserving the some of the distinctive characteristics of that group.

Category	Pronouns
We	We, Our, Us
You	You, Your, Yourself
Me	Me, My, Myself, I
They	They, Them, Their
He	He, Him, Himself, His
She	She, Her, Herself, Hers
It	It, Itself, Oneself

Table 2: Pronouns divided into 7 categories and used to quantify fraction features.

Next, considering the unbalanced distribution in the data (see Figure 5), a simple linear regression formulation will have a bias towards the most frequent rigidity scores. To overcome this challenge, we used a weighted least squares (WLS) model. WLS models are effective in dealing with heteroscedasticity (i.e., non-constant variance in the error term). Using WLS, the closed form solution for the regression problem is

$$\hat{\beta}_{WLS} = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{Y} \tag{3}$$

such that W is the weighting matrix and it is defined as

$$\mathbf{W} = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix} \tag{4}$$

where w_i is weight corresponding to document i. In this paper, we choose to set the document's weight to the reciprocal of each value-motivated group's variance, $\sigma^2_{G(i)}$ where $G(i)$ is the value-motivate group that authored document i. Finally, groups might have zero variance, and therefore, we added to small constant ($\epsilon = 10^{-5}$) to avoid undefined weights,

$$w_i = \frac{1}{\sigma^2_{G(i)} + \epsilon} \tag{5}$$

Using this approach, we can give higher weight to documents authored by groups with lower variance.

5. Experimental Design and Evaluation

We performed a number of experiments to analyze the quality of features and predictive models. First, we quantified the 17 features, mentioned above, on all 342 documents. Next, we analyzed changes in feature values for the WBC between 2008 and 2016. In the previous example in Figure 4, we observed a change in the rigidity score after the church won the supreme court case. We were interested in observing any changes in feature values following that event. Figure 6 shows two of the features, namely fraction of Green et al.'s judgment sentences and fraction of unique words. We observe a sudden drop in WBC's judgments and an increase in the percentage of unique words. Then, both features continued to follow their overall trends (increasing fraction of judgment and decreasing fraction of unique words). This example highlights the importance of single document analysis which would provide real time insights into major events that would change a group's ideology and/or behaviour. Such events can be known to the public such as a court ruling or unknown such as a change in leadership or loyalty, an inner-group division, etc. Single document analysis can revel such changes which will lead to adjustment decision making towards value-motivate groups (e.g., starting negotiations).

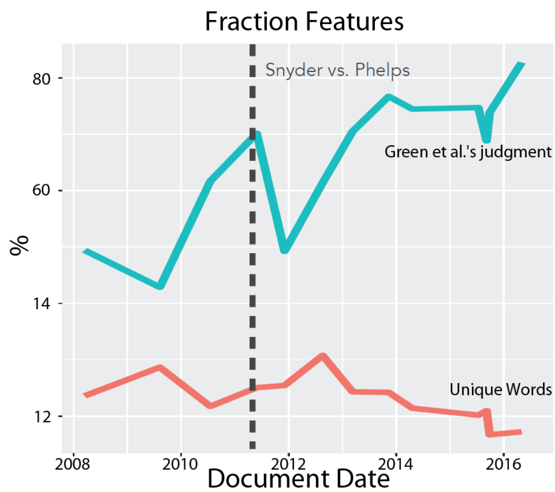


Figure 6: Fraction of Green et al.'s judgment and unique words in single documents from the Westboro Baptist Church between 2008 and 2016. Changes in the feature values are observed following the Snyder v. Phelps ruling.

Next, we compared the feature values of group-based and individual-based documents. Green et al. found interesting correlations between the linguistic rigidity score of value-motivated groups and some of the fraction predictors, namely pronoun + value word judgments, "We" pronoun, and "They" pronoun (Green et al. 2017). We performed the same analysis using single documents (see Figure 7). We observe a significant change in the distribution of feature values, especially the judgment and "We" fraction features. This indicates that these features were capturing the writing style of groups, and therefore, having a significant contribution to the process of predicting their linguistic rigidity. As single document within the same group varies in their rigidity score, the previous features fail in relating the writing style to the outcome.

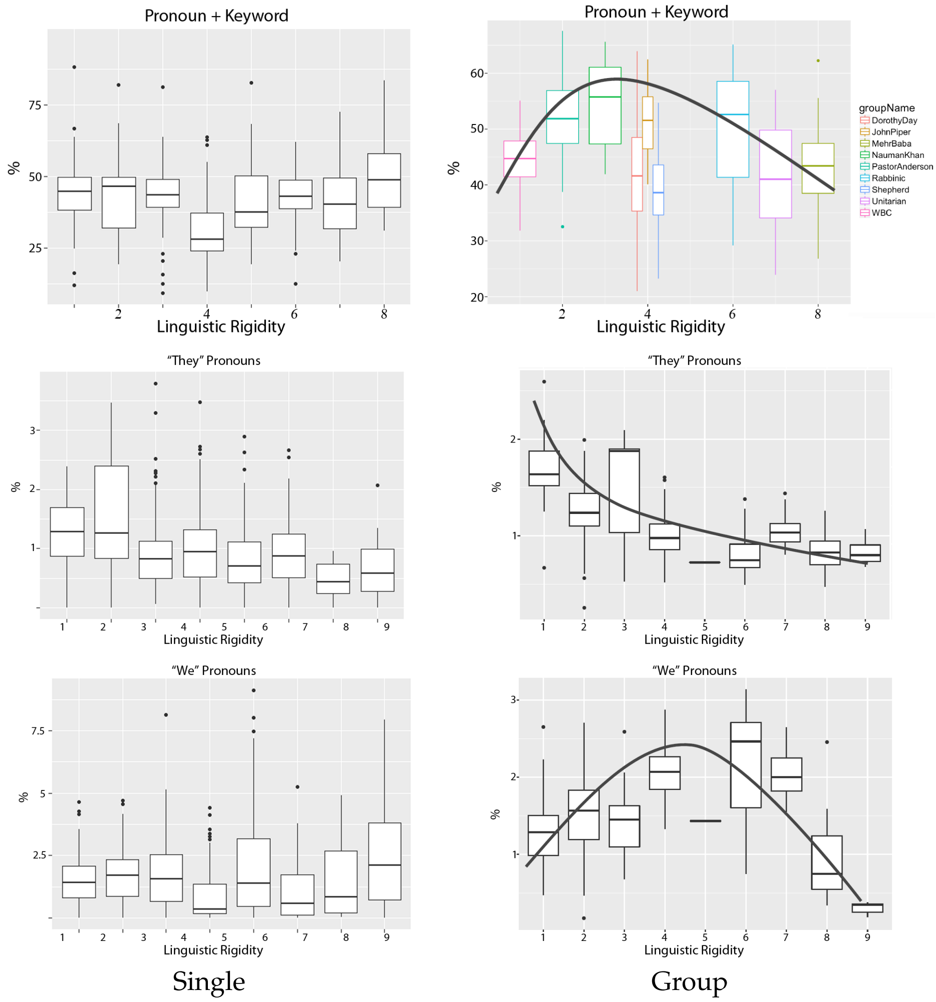


Figure 7: Comparing fraction feature values (namely, Green et al.'s judgment, "We" and "They" pronouns) between value-motivated groups and single documents.

To better understand the relationship between predictor variables and the response, we performed a correlation analysis using a linear regression model. Table 3 shows the coefficients of the linear regression model that correspond to each to the features. From these results, we observer that 1) none of the sentiment, value words and judgment feature are statistically significant in capturing single documents' rigidity, 2) three out of seven pronoun fraction features are significant, and these features represent authors talking about themselves or their value-motivated group (e.g., I, my, we, it, our, etc.) Most importantly, the coefficients of these pronouns are positive meaning that the more the authors talk about themselves, the higher is their rigidity score and the more flexible is their text. Finally, we observe that most of the group names have significant coefficients. This indicates that knowing the value-motivated group of a single document is

Table 3: Correlation analysis of linguistic rigidity using linear regression.

Dependent variable: ***Linguistic Rigidity Score***			
Predictor	**Estimate (Std. Error)**	**Predictor**	**Estimate (Std. Error)**
perPos	−0.058 (0.491)	uniqueWordsFrac	−0.565 (0.674)
perNeg	−0.602 (0.541)	group.AEU	4.635^{***} (0.478)
posDoc	0.212 (0.183)	group.Anderson	-2.763^{***} (0.675)
avgSD	0.811 (0.679)	group.Bahai	1.654^{***} (0.379)
avgEVC	−0.355 (0.352)	group.DorothyDay	0.284 (0.434)
vJudgementCount	−0.430 (1.143)	group.ISIS	-1.714^{***} (0.454)
vJudgementFrac	−0.215 (0.606)	group.JohnPiper	−0.117 (0.429)
gJudgementFrac	−0.890 (0.603)	group.LiberalJudaism	−0.433 (0.388)
weFrac	1.108^{**} (0.495)	group.MalcolmX	-1.578^{***} (0.500)
youFrac	−0.256 (0.472)	group.MehrBaba	2.151^{***} (0.375)
meFrac	2.031^{***} (0.650)	group.Rabbinic	0.125 (0.406)
theyFrac	0.128 (0.371)	group.Shepherd	-1.094^{***} (0.404)
heFrac	0.748 (0.461)	group.Unitarian	0.652 (0.499)
sheFrac	0.321 (0.750)	group.WBC	-2.264^{***} (0.457)
itFrac	1.098^{*} (0.632)	group.YV	1.620^{***} (0.420)
Intercept	2.954^{***} (0.978)		
Observations	342		
R^2	0.693		
Adjusted R^2	0.663		
Residual Std. Error	1.052 (df = 311)		
F Statistic	23.362^{***} (df = 30; 311)		
Note:			*p<0.1; **p<0.05; ***p<0.01

very important for estimating its rigidity score. We further observe that the coefficient of both sides of the rigidity scale are consistent with their score, i.e., ISIS, Pastor Anderson, and WBC have negative coefficients while Unitarian, American Ethical Union (AEU), and Meher Baba have positive coefficients.

Method	Training MAE	Training ACC(%)	10-Fold CV Testing MAE	10-Fold CV Testing ACC(%)
Least Squares LR	0.7127	52.34	0.7975	48.87
Weighted Least Squares LR	0.7198	54.39	**0.7485**	**52.10**
Regression Random Forest	**0.3465**	**71.35**	0.8458	44.70
Regression Boosted Tree	0.5873	40.06	1.1621	21.64
Support Vector Regression Machine	0.6905	54.97	0.7742	50.62

Table 4: Predictive performance of our models on single documents

Next, we use partial F-Test (Jamshidian, Jennrich, and Liu 2007) to test the hypothesis that a model including the group name feature is significantly better than a model excluding it. The test obtains a F-value of 32.817 and a P-value < 2.2e-16. These results confirm the hypothesis and indicates that the group name feature is significantly essential from estimating the linguistic rigidity score of single documents.

Finally, we evaluated the predictive capability of our models in terms of the mean absolute error (MAE) and classification accuracy (ACC). We calculated the ACC by rounding the estimated rigidity scores to the nearest integer. The motivation for calculating the ACC is to analyze the per-class accuracy and build a confusion matrix. Also, along with the least squares (LS) and the WLS regression models, we fitted a regression random forest model (Liaw, Wiener et al. 2002), a support vector regression machine model (Drucker et al. 1997), and a regression boosted tree model (Elith, Leathwick, and Hastie 2008). Table 4 shows the 10-folds cross validation testing performance. The table also includes the training error (i.e., the MAE and the ACC on a model trained and tested on all 342 documents). We used the training error as an indicator of how could our models can perform considering the unbalanced nature of the data. From the performance results, we observe the following: 1) random forest achieved the best training performance, but rather a worse testing performance. This indicates that random forest, when applied on unbalanced data, will overfit the training observations as it has a lower bias and higher variance; 2) using a weighted least squares solution for the linear regression model improved the training and testing performance significantly. Adding the weights further help in reaching a testing classification accuracy as close as the training one for least squares (52.1% using WLS compared to 52.34% using LS).

To further understand the usefulness of adding the weights, we visualized the 10-folds averaged confusion matrices for LS and WLS (see Figure 8). We observed that adding the weights helped in modeling documents at both ends of the linguistic rigidity scale (i.e., documents of scores 1,2,6, and 8). These documents together constitute only 37.71% of all documents. A standard LS model underperformed on these documents as it is biased towards documents of most frequent rigidity scores, specifically 3 and 4. Also, both LS and WLS models failed in correctly classifying any of the documents with a rigidity score of 7. Nonetheless, WLS models were able to classify 18% more of these documents as 6 rather than 5, and therefore, reducing the overall absolute error.

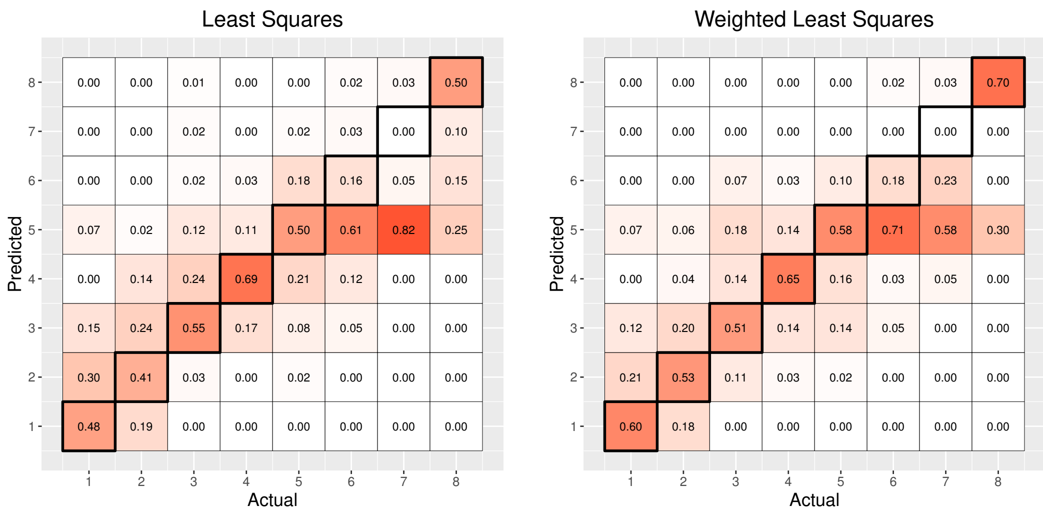


Figure 8: Confusion matrices of least squares and weighted least squares linear regression models.

6. Conclusions

In this work, we extended a computational linguistic analysis proposed by Green et al. to individual documents. Although the linguistic rigidity analysis can provide insights into the behaviour of value-motivated groups, it fails short in assisting real time decision making. The main contribution of this work includes, 1) analyzing the importance of existing features for predicting single documents rigidity scores, 2) adding an additional categorical feature that quantifies the name of the value-motivated group, and 3) proposing a weighting scheme and using it to fit weighted least squares linear regression models. The experimental results indicates that: 1) previous features, which were effective in estimating groups' scores, have insignificant correlation with the rigidity scores of single documents. This could indicates that these features were capturing the writing style rather than the rigidity score of value-motivated groups; 2) using a correlation analysis and a partial F-test, we found that the newly added group name feature has a significant contribution in prediction process; and 3) weighting documents by the inverse of their group variance significantly improved the prediction performance especially for documents of infrequent scores at both ends of the linguistic rigidity scale. This work can be extended in several directions. We would like to refine the feature set by: 1) using feature selection methods such as Chi-square and information gain (Yang and Pedersen 1997) to select better value words, and 2) exploring the usage of words embedding to calculate distances between judgment sentence (Kusner et al. 2015), and therefore, construct semantic density features from these sentences. Most importantly, we would like to investigate the use of our models for predicting groups' behaviour. We will analyze text documents produced by value-motivated groups[2], and use their predicted rigidity scores as a predictor for changes in their behaviour (i.e., becoming violent and/or being added to terrorist lists).

References

Bird, Steven, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Drucker, Harris, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. 1997. Support vector regression machines. In *Advances in neural information processing systems*, pages 155–161.

Elith, Jane, John R Leathwick, and Trevor Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4):802–813.

Estrada, Ernesto and Juan A Rodriguez-Velazquez. 2005. Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103.

Green, Seth, Megan Stiles, Katherine Harton, Samantha Garofalo, and Donald E Brown. 2017. Computational analysis of religious and ideological linguistic behavior. In *Systems and Information Engineering Design Symposium (SIEDS), 2017*, pages 359–364, IEEE.

Hu, Minqing and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177, ACM.

Jamshidian, Mortaza, Robert I Jennrich, and Wei Liu. 2007. A study of partial f tests for multiple linear regression models. *Computational Statistics & Data Analysis*, 51(12):6269–6284.

Kröll, Mark and Markus Strohmaier. 2009. Analyzing human intentions in natural language text. In *Proceedings of the fifth international conference on Knowledge capture*, pages 197–198, ACM.

Kusner, Matt, Yu Sun, Nicholas Kolkin, and Kilian Weinberger. 2015. From word embeddings to document distances. In *International Conference on Machine Learning*, pages 957–966.

2 Foreign Terrorist Organizations https://www.state.gov/j/ct/rls/other/des/123085.htm

Landrum, Nancy E, Connor Tomaka, and John McCarthy. 2016. Analyzing the religious war of words over climate change. *Journal of Macromarketing*, 36(4):471–482.

Lewis, David D, Yiming Yang, Tony G Rose, and Fan Li. 2004. Smart stopword list.

Liaw, Andy, Matthew Wiener, et al. 2002. Classification and regression by randomforest. *R news*, 2(3):18–22.

Manning, Christopher D, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to information retrieval*, volume 1. Cambridge university press Cambridge.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

Miller, Kent D. 2002. Competitive strategies of religious organizations. *Strategic Management Journal*, 23(5):435–456.

Pennington, Jeffrey, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *EMNLP*, volume 14, pages 1532–1543.

Sagi, Eyal, Stefan Kaufmann, and Brady Clark. 2009. Semantic density analysis: Comparing word meaning across time and phonetic space. In *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics*, pages 104–111, Association for Computational Linguistics.

Venuti, Nicholas, Brian Sachtjen, Hope McIntyre, Chetan Mishra, Matthew Hays, and Donald E Brown. 2016. Predicting the tolerance level of religious discourse through computational linguistics. In *Systems and Information Engineering Design Symposium (SIEDS), 2016 IEEE*, pages 309–314, IEEE.

Wade-Benzoni, Kimberly A, Andrew J Hoffman, Leigh L Thompson, Don A Moore, James J Gillespie, and Max H Bazerman. 2002. Barriers to resolution in ideologically based negotiations: The role of values and institutions. *Academy of Management Review*, 27(1):41–57.

Willett, Peter. 2006. The porter stemming algorithm: then and now. *Program*, 40(3):219–223.

Yang, Yiming and Jan O Pedersen. 1997. A comparative study on feature selection in text categorization. In *ICML*, volume 97, pages 412–420.