

# Simulating Audiences

## Automating Analysis of Values, Attitudes, and Sentiment

Thomas Clay Templeton, Kenneth R. Fleischmann, and Jordan Boyd-Graber

College of Information Studies

University of Maryland

College Park, MD

{clayt,kfleisch,ying}@umd.edu

**Abstract**—Current events such as the Park51 Project in downtown Manhattan create “critical discourse moments,” explosions of discourse around a topic that can be exploited for data gathering. Policymakers have a need to understand the dynamics of public discussion in real time. Human values, which are cognitively related to attitudes and serve as reference points in moral argument, are important indicators of what’s at stake in a public controversy. This work shows that it is possible to link values data with reader behavior to infer values implicit in a topical corpus, and that it is possible to automate this process using machine learning. Thus, this work represents a preliminary attempt at scalable computational social science using techniques from natural language processing.

Keywords—human values; natural language processing; crowdsourcing

### I. INTRODUCTION: VALUES IN THE READING PROCESS

Sentiment analysis allows the automatic detection of opinion-bearing text. The ability to quickly and efficiently gauge a broad range of reactions to policies, products, and people has been quickly adopted and developed by academics, industry, and government [1]. However, the information provided by conventional sentiment analysis is limited because it ignores the motivations behind sentiment [2].

In this paper, we explore the role of human values in online discourse. Previous work on scaling up values analysis has drawn on classic content analysis techniques [3]: annotators manually label text that expresses a value and computers try to mimic the annotation process on previously unlabeled text. Although this approach is standard in NLP [4], it transfers unresolved challenges from the manual annotation of values into automated values analysis. The correct mapping of *text* into an abstract map of human values is a subtle problem for humans to agree on, especially when they identify differently with different values. Annotating values in text lacks a tradition of synthesis and standardization, and so recent research has focused on filling that gap [5].

In contrast, we focus on the assessment of the values of an *individual*, rather than a span of text. Assessing the values of individuals benefits from a tradition of synthesis and standardization based on studying human values via surveys [6]. Below, we describe the theoretical basis for how we use the Portrait Values Questionnaire (PVQ) [7] to measure the

values of individuals, and then relate those values with an underlying property of the text, sentiment, to predict the behavior of a reader, such as whether they agree or disagree with the text that they are reading, in order to understand the relationship between the values of readers and their attitudes toward current events.

The classic linear communication paradigm [8] can be summarized as follows:

“Who...Says What...In Which Channel...To Whom... With What Effect?”

Shannon and Weaver’s classic treatment reflects the same basic functional decomposition [9]. Berlo [10] informally adapts Shannon and Weaver’s model to human/ human communication, considering personal characteristics as factors in the encoding and decoding functions. Today, the Natural Language Processing (NLP) community commonly models aspects of the generation of language as “noisy channel” processes [11].

In this work, we model the reading process as a probabilistic process (see Fig. 1) through which the outcome is influenced by the properties of the text (*t*) and the state of the reader (*r*). The end result is a behavior by the reader (*b*).

In other research contexts, outcome behaviors might include replying to an email, clicking the “like” button on Facebook, or retweeting. Similarly, properties of state might include demographic variables, network variables, or personal history.

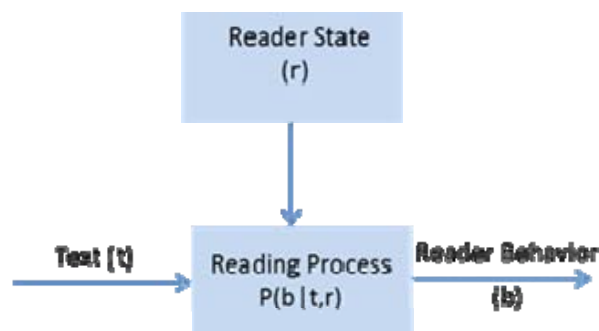


Figure 1. Formal model of the reading process

In this work, reader state ranges over a set of human values developed by Shalom Schwartz. Humans share a universal map of values [12], Schwartz argues, but the weights on particular values vary. A social group’s weighted value set is the backdrop against which moral argumentation unfolds. Proposals concerning the value of an entity or a course of action tend to refer back, at least implicitly, to fundamental moral premises in the abstract space of human values [13, 14].

A couple of toy examples:

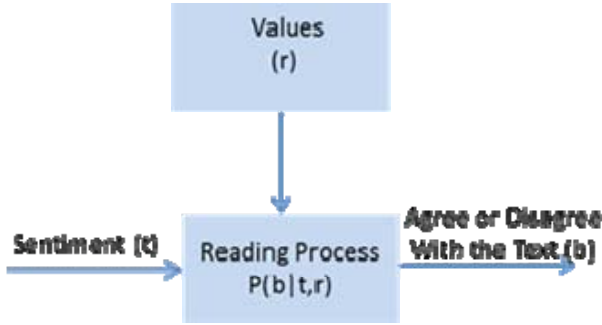
We all agree that safety is good. Installing traffic lights will promote safety. Therefore, we should install traffic lights.

All our stakeholders want wealth. Investing is likely to bring us wealth. Therefore, we should invest.

As shared reference points in value-laden arguments, values are an important variable in attitude formation and attitude change [15, 16]. In public persuasion campaigns, the values of the audience are a key consideration in crafting a message [17, 18].

In the experiments that follow, we consider how an audience’s values are related to a simple behavioral outcome. After exposing readers to texts about an issue in the news, we ask them to quantify their response on a scale ranging from strong disagreement to strong agreement.

Based on crowdsourced sentiment annotation, each text is assigned a score representing the attitude it expresses about the issue. The operationalized model informing the experiments is summarized in Fig. 2. Binarizing the reader state variable will later facilitate accessible visualization and analysis of the results.



In Fig. 2, we apply the approach from Fig. 1 to this case, such that values serve as the reader state ( $r$ ), which influences the reader behavior ( $b$ ) to agree or disagree with the sentiment expressed in the text ( $t$ ).

## II. METHODS

### A. Data Collection

Amazon’s Mechanical Turk is a marketplace that connects “Requestors” willing to pay people to complete simple tasks with “Turkers” willing to do the tasks. Mechanical Turk enables requestors to post tasks, which Amazon refers to as Human Intelligence Tasks or “HITs.”

We used Mechanical Turk’s qualification process to require readers to complete a 21-question version of the Schwartz Portrait Values Questionnaire (PVQ), along with some demographic questions, before proceeding to our HITs. Each HIT consisted of a series of opinionated paragraphs along with prompts for Turkers to quantify their response.

For this study, we focused on a particular controversy, the Park51 controversy, also referred to as the “Downtown Manhattan Islamic Community Center” or the “Ground Zero Mosque.” This controversy involves strong polarization of sentiment and a large number of opinionated texts directly related to this controversy were available via Google News.

In the first experiment, we asked 53 Turkers to rate 50 opinionated news and opinion paragraphs about the Park51 controversy and then to write a paragraph describing their perspective on the Park51 controversy.

In the second experiment, we presented the 53 paragraphs collected in the first experiment to 100 Turkers. We asked Turkers to indicate agreement or disagreement with each paragraph using a scale from 1 (I strongly disagree) to 4 (I strongly agree).

In the third experiment, we presented the same 53 paragraphs to a different set of 100 Turkers to evaluate sentiment expressed in every paragraph using a scale from 1 (The author is strongly against the project) to 4 (The author is strongly in favor of the project).

### B. Data Analysis

After performing the experiments just described, we binarized the data for each Schwartz value, discarding data points where a reader’s PVQ score was equal to the median PVQ score for that value. From each of the two reader groups thus obtained, we took an average reader response score for each paragraph.

To test whether the values of a reader can help explain their agreement with a piece of text, we ran a regression model of the form:

$$\vec{A} = \beta_0 + \beta_1 \vec{S} + \beta_2 \vec{V} + \beta_3 \vec{S} \circ \vec{V} \quad (1)$$

In Eq. 1,  $V$  indicates whether the data point belongs to the high or low value level,  $S$  is the average sentiment of a paragraph, and  $A$  is the average reader agreement for that paragraph, for that value level (for 106 total data points). The coefficients are fit using ordinary least squares; if the coefficients associated with the values are zero, then there is no explanatory power from including the values of an individual.

We then performed a second set of regressions of the same form (Eq. 1), using average agreement scores predicted by SVM regression [19] for  $A$ . We used unigrams weighted by tf-idf to train the SVMs. A high value and low value SVM was trained for each of the 10 values, for a total of 20 sets of SVM regressions. A leave-one-paragraph-out approach was employed for all 20 sets of SVM regressions.

For each instantiation of Eq. 1 using the SVM predictions for A, the high value and the low value classifiers each provided half the data points, one for each paragraph, for a total of 106 data points in the regression for each value.

### III. RESULTS AND DISCUSSION

Table 1 shows the parameter estimates and standard errors for three of the four parameters in the regression models for each value. The intercept is not shown, since it is not relevant to the relationship between agreement, sentiment, and values. For each value, models trained using actual and predicted agreement scores, respectively, are shown.

TABLE I. RESULTS FROM REGRESSION MODEL

Model	Value Parameter	Sentiment Parameter	Interaction Parameter
Benevolence Actual	.03 (.12)	.249*** (.030)	.045 (.042)
Benevolence Predicted	-.934 (.84)	.11 (.84)	.23 (.30)
Universalism Actual	-1.86*** (.12)	-.131*** (.030)	.721*** (.043)
Universalism Predicted	-4.9 *** (1.3)	-.60 (.32)	1.46** (.45)
Self Direction Actual	.48*** (.12)	.338*** (.030)	-.168*** (.042)
Self Direction Predicted	1.10* (.81)	.33 (.20)	-.28 (.29)
Stimulation Actual	-.15 (.11)	.185*** (.027)	.108** (.039)
Stimulation Predicted	-.58 (.71)	.05 (.18)	.22 (.25)
Hedonism Actual	.02 (.12)	.268*** (.029)	.002 (.041)
Hedonism Predicted	-.09 (.75)	.15 (.19)	.10 (.266)
Achievement Actual	.84*** (.12)	.41 *** (.030)	-.26*** (.043)
Achievement Predicted	1.61 (.93)	.48* (.23)	-.50 (.33)
Power Actual	1.37*** (.11)	.53*** (.028)	-.48*** (.040)
Power Predicted	2.8* (1.1)	.74* (.28)	-0.79* (.40)
Security Actual	2.79*** (.13)	.585*** (.033)	-1.038*** (.047)
Security Predicted	7.3 *** (1.6)	.87*** (.41)	-2.20*** (.58)
Conformity Actual	1.07*** (.11)	.454*** (.028)	-.362*** (.040)
Conformity Predicted	2.28* (.95)	.55* (.24)	-.60 (.34)
Tradition Actual	0.85*** (.13)	.364*** (.034)	-.270*** (.048)
Tradition Predicted	1.74* (.83)	.29 (.30)	-.52 (.30)

Remarkably, the regressions on predicted agreement scores reproduced the correct sign for every significant parameter in the regressions on actual agreement scores. Further, there are no false positives in the data from predicted scores; each significant result from the predicted data is at least as significant for the actual data. This suggests that our approach is promising for identifying values that are related to the reception of text. Two parameters are especially key to this analysis: the value parameter and the interaction parameter.

#### A. Value parameter

The value parameter indicates the change in intercept when the value level changes from 0 (low) to 1 (high). A positive parameter for this term indicates that people holding the value tend to agree with anti-Park51 paragraphs more than people who do not hold the value.

Regressions on the following values produced significant positive value parameters. That is, a person holding any one of these values is more likely to agree with an anti-Park51 paragraph than a person not holding the value:

Self Direction\*\*\*, Achievement\*\*\*, Power\*\*\*, Security\*\*\*, Conformity\*\*\*, Tradition\*\*\*

Regressions on the following values produced significant negative value parameters. That is, a person holding this value is more likely to disagree with an anti-Park51 paragraph than a person not holding the value:

Universalism\*\*\*

The regressions based on the predicted agreement scores reproduced the correct sign for all of these parameters, although the value parameter was only significant to  $p < 0.001$  for 2 of these (Security and Universalism) and to  $p < 0.05$  for an additional four (Self-Direction, Power, Conformity, Tradition).

#### B. Interaction parameter (Value\*Sentiment)

The parameter associated with the interaction term indicates the difference in slope of the regression lines relating sentiment and agreement between the case where the value is 0 (low) and the case where it is 1 (high). A positive interaction parameter means that people who hold the value react more positively to incremental increases (toward positivity) of paragraph sentiment than do people who do not hold the value.

Regressions on the following values produced a positive interaction parameter, meaning the change in Agreement associated with moving in the pro-Park51 direction along the Sentiment axis was greater for readers who held the value than for readers who did not:

Universalism\*\*\*, Stimulation\*\*

Regressions on the following values produced a negative interaction parameter, meaning the change in Agreement associated with moving in the pro-Park51 direction along the Sentiment axis was less for readers who held the value than for readers who did not:

Self-Direction\*\*\*, Achievement\*\*\*, Power\*\*\*, Security\*\*\*, Conformity\*\*\*, Tradition\*\*\*

The regressions based on the predicted agreement scores reproduced the correct sign for all of these parameters, although the interaction parameter was only significant to  $p < 0.001$  for 2 of these (Security and Universalism) and to  $p < 0.05$  for an additional one (Power).

All of the values listed above except one (Stimulation) follow a common pattern: the sign of the interaction term is opposite the sign of the values term. In other words, the regression lines move toward convergence (and possible intersect) from the intercept moving to the right in the graph.

As in another analysis of a similar corpus, security and universalism appear to be the most salient values for this issue [20].

#### IV. LIMITATIONS AND FUTURE DIRECTIONS

One limitation of this study is that it focuses on an issue (Park51) on which people appear to be rigidly divided into two groups; people tend to agree either with most texts in support of or with most texts against the proposed project. The sentiment expressed in text appears to play a very large role in deciding a particular reader's response to the text.

In the face of a more nuanced issue like nuclear power, our approach breaks down, since the salient features of the text cannot as faithfully be boiled down to points along a sentiment axis [21]. Rather, in nuclear discourse, the framing of a piece of text, over and above the sentiment it expresses, seems to matter [22]; moreover, these framings may be related to values [23]. Perhaps some combination of manual thematic analysis on the corpus and factor analysis on values and reader responses can untangle the web of relevant textual features.

#### V. CONCLUSION

In this work, we demonstrated a technique for automatically assessing which Schwartz values align with each side in a public debate. Additionally, we demonstrated that SVM regression can be used to simulate audiences defined by their affinity for a value or lack thereof. We used these SVM regressions to predict average agreement, and showed that regressions using the predicted scores performed comparably to regressions using actual average value scores in identifying which values aligned with which side in the debate.

Values analysis can be an important complement to sentiment analysis, as values give insight into why people hold particular opinions. At the population and corpus level, values indicate what considerations are important for different sides of a public policy debate. Scaling up values analysis within and across topics can provide timely information to social scientists and policymakers about the relevant dimensions of public opinion about controversial topics.

#### REFERENCES

- [1] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, pp. 1-135, 2008.
- [2] S. Kim and E. Hovy, "Automatic identification of pro and con reasons in online reviews," in *Proceedings of the COLING/ACL 2006*, pp. 483-490, 2006.
- [3] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications, 2003.
- [4] R. Artstein and M. Poesio, "Inter-coder agreement for computational linguistics," *Computational Linguistics*, vol. 34, pp. 555-596, 1992.
- [5] A.-S. Cheng and K. R. Fleischmann, "Developing a meta-inventory of human values," in *Proceedings of the 73rd Annual Meeting of the American Society for Information Science and Technology*, 2010.
- [6] V. A. Braithwaite and W. A. Scott, "Values," in *Measures of Personality and Social Psychological Attitudes*, J. P. Robinson, P. R. Shaver, and L. S. Wrightsman, Eds., Burlington, MA: Academic Press, 1991, pp. 661-776.
- [7] S. H. Schwartz, "Value orientations: Measurement, antecedents, and consequences across nations, in *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, R. Jowell, C. Roberts, R. Fitzgerald, & G. Eva, Eds., Thousand Oaks, CA: Sage, 2007, pp. 169-203.
- [8] H. Lasswell, "The structure and function of communication in society," in *The Communication of Ideas*, L. Bryson, L., Ed., New York: Harper and Row, 1948.
- [9] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1949.
- [10] J. Berlo, *The Process of Communication: An Introduction to Theory and Practice*. New York, NY: Holt, Rinehart, and Winston, 1963.
- [11] C. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*. Cambridge: MIT Press, 1999.
- [12] S.H. Schwartz, "Are there universal aspects in the structure and content of human values?," *Journal of Social Issues*, vol. 50, pp. 19-45, 1994.
- [13] W. R. Fisher, "Toward a logic of good reasons," *Quarterly Journal of Speech*, vol. 64, pp. 376-384, 1978.
- [14] K. R. Wallace, "The substance of rhetoric: Good reasons," in W. L. Benoit, D. Hampe, and P. J. Benoit, Eds., *Readings in Argumentation*, Netherlands: Walter de Gruyter & Co, 2002, pp. 383-398.
- [15] S. Hitlin and J. A. Piliavin, "Values: Reviving a dormant concept," *Annual Review of Sociology*, vol. 30, pp. 359-393, 2004.
- [16] M. Rokeach, *The Nature of Human Values*. New York: Free Press, 1973.
- [17] R. B. Cialdini, *Influence: The Psychology of Persuasion*. New York: Quill, 1993.
- [18] W. K. McGuire, "Input and output variables currently promising for constructing persuasive communications," in *Public Communication Campaigns*, R. E. Rice, C. K. Atkin, Eds., Thousand Oaks, CA: Sage Publications, 2001, pp. 22-48.
- [19] T. Joachims, "Making large-Scale SVM Learning Practical," in *Advances in Kernel Methods - Support Vector Learning*, B. Schölkopf and C. Burges and A. Smola, Eds., Cambridge, MA: MIT Press, 1999.
- [20] T. C. Templeton, K. R. Fleischmann, and J. Boyd-Graber, "Comparing values and sentiment using Mechanical Turk," in *Proceedings of the 2011 iConference*, 2011.
- [21] T.C. Templeton, K. R. Fleischmann, and J. Boyd-Graber, "The Relationship between Human Values and Attitudes Toward the Park51 and Nuclear Controversies," in *Proceedings of the American Society for Information Science and Technology*, 2011.
- [22] W.A. Gamson and A. Mogdighiani, "Media Discourse and Public Opinion," *The American Journal of Sociology*, vol. 95, pp. 1-37.
- [23] D. Chong and J.N. Druckman, "Framing Theory," *Annual Review of Political Science*, vol. 10, pp. 103-126.