Don't Just Ask Me for Facts!
Measuring Political Sophistication with Open-Ended Responses

Laura Buchanan[*]   Patrick Kraft[†]

April 24, 2016

Abstract

...

Keywords: ...

---

[*]lcb402@nyu.edu
[†]pk1622@nyu.edu

1 Introduction

On of the most fundamental concepts in the study of political attitudes and behavior is political sophistication and knowledge (Converse, 1964; Delli Carpini and Keeter, 1996). However, its measurement has been subject to re-occurring scholarly debate (e.g. Mondak, 2000; Mondak and Davis, 2001; DeBell, 2013). Many analyses exclusively rely on individual levels of political information measured by factual knowledge questions. However, recent research points to important differences between types of knowledge questions that have previously been disregarded (Barabas et al., 2014). Furthermore, scholars argued that the conceptualization of political sophistication should take into account how people structure their attitudes and beliefs, rather than just focusing on pure levels of information (e.g. Luskin, 1987). As such, measuring sophistication solely based on answers to political trivia may misclassify respondents who cannot recall these facts, but do indeed have a coherent network of political ideas.

The paper proposed here examines alternative measures of political sophistication based on individual responses to open-ended likes and dislikes questions in the 2012 American National Election Survey (ANES). For this portion of the survey, respondents write freely about what they like and dislike about the Republican and Democratic parties and presidential candidates. We leverage recent advances in text analysis to make inferences about the respondents' levels of political sophistication and belief constraint by focusing on how they describe their attitudes towards different political actors. More specifically, we consider the lexical complexity and readability of responses, as well as the diversity in topics raised by respondents based on structural topic models (Roberts et al., 2014), in order to assess the degree to which political attitudes are structured and expressed in a more complex manner. We suspect that the diversity in topics a respondent mentions in their answers, or the detail with which they speak about the topics they mention, will covary with the other political knowledge measures available in the data. We therefore compare the text-based measures to common factual knowledge items as well as the interviewer assessment of the respondent's

political knowledge as benchmarks.

All responses from a single participant will be considered a single document. First, we will find the topics present over the entire corpus, starting with a simple off-the-shelf baseline model, and improve performance as we learn the knobs of topic models as they are covered in class. Then, for each document, we will measure the diversity in estimated topic proportions. This will be our first measure. Second, we will look at the length of the document. Third, we will look at document length relative to the number of topics. We suspect that the first measure will correlate with political knowledge, displaying a wide spectrum of political insight. The second measure will act like something of a control or baseline; can we simply predict political knowledge by the length of response? Finally, the third measure may help us gain insight into potential cases where a respondent focuses on very few topics, but has a wealth of insight into the topics they choose to mention.

The dataset is available from American National Election Studies. For his PhD, Patrick has cleaned this data to some degree, using an implementation of the Aspell spell checking algorithm in R to correct for misspellings. The dataset may be further cleaned in two different ways. First, we may account for eloquence of the respondents, using readability or type-to-token ratio of the responses. We would like to capture political intelligence specifically, rather than general intelligence. Similarly, we may correct for education level, which is also available in the data. Second, we may create a list of "talking points" that could possibly turn up in the responses. We suspect that specific phrasing could suggest that a respondent is parroting a political message, rather than comprehending and reframing the message in their own words. This may be too difficult a task to complete for this project, but a good first pass try would include finding common n-grams, up to n = 6, and use a web search to determine whether these are known slogans. Respondents would then be "punished" if they primarily use these slogans, as that suggests that they do not have a deep political knowledge.

To assess the performance of our model, we will compare our results to the other political

knowledge scores, perhaps using a method such as the Adjusted Rand Index. An evaluation such as this will be useful, because we can check, first, if our method is performing better than chance, and, if so, is it performing poorly, well, or excellently. However, we will also take a qualitative look at the performance our model, to determine if we are able to find cases of political knowledge not captured in the standard measure of political knowledge.

Overall, we hope to gain important insights about individual political sophistication conceptualized as the structure and constraint of political beliefs (for example, in line with Tetlock's 1983 notion of integrative complexity) by taking into account open-ended survey responses. Furthermore, developing valid measures of political sophistication based on open-ended responses will provide new opportunities for comparisons of political knowledge across time and contexts.

2 Theoretical Framework

2.1 Political Knowledge and Sophistication

2.2 Open-Ended Responses

2.3 Hypotheses

3 Empirical Results

3.1 Data Overview

3.2 Specification of Topic Model

3.3 Performance Evaluation

4 Conclusion

References

Barabas, Jason, Jennifer Jerit, William Pollock, and Carlisle Rainey. 2014. "The Question(s) of Political Knowledge." *American Political Science Review* 108 (04): 840–855.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: Free Press.

DeBell, Matthew. 2013. "Harder than it looks: Coding political knowledge on the ANES." *Political Analysis* 21 (4): 393–406.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, London: Yale University Press.

Luskin, Robert C. 1987. "Measuring political sophistication." *American Journal of Political Science*: 856–899.

Mondak, Jeffery J. 2000. "Reconsidering the measurement of political knowledge." *Political Analysis* 8 (1): 57–82.

Mondak, Jeffery J, and Belinda Creel Davis. 2001. "Asked and answered: Knowledge levels when we will not take 'don't know' for an answer." *Political Behavior* 23 (3): 199–224.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58 (4): 1064–1082.

Tetlock, Philip E. 1983. "Cognitive style and political ideology." *Journal of Personality and Social Psychology* 45 (1): 118.