# Looking for Answers

## A Naive Approach for Measuring Political Sophistication

Patrick W. Kraft

113th APSA Annual Meeting

San Francisco, 1 September 2017

Stony Brook University

# Who scores higher on political knowledge?

| | **Respondent A** | **Respondent B** |
|---|---|---|
| Obama (+) | | |
| Obama (-) | | |
| Romney (+) | | |
| Romney (-) | | |

# Who scores higher on political knowledge?

| | **Respondent A** | **Respondent B** |
|---|---|---|
| Obama (+) | I think he is honest, has good intentions. | |
| Obama (-) | I don't feel he is up for the job, he doesn't really know how to get things accomplished from idea to actual reality. | |
| Romney (+) | He comes across as an honest person and I feel that financially he would be better for the country. | |
| Romney (-) | I am a moderate conservative and there are some things about anti-gay rights that I don't support. | |

# Who scores higher on political knowledge?

| | **Respondent A** | **Respondent B** |
|---|---|---|
| Obama (+) | I think he is honest, has good intentions. | He's got a hot wife. Sound mind. Bill Clinton, the best president we ever had, is voting for him. He's not in any type of scandal and is the perfect president for us now. |
| Obama (-) | I don't feel he is up for the job, he doesn't really know how to get things accomplished from idea to actual reality. | |
| Romney (+) | He comes across as an honest person and I feel that financially he would be better for the country. | |
| Romney (-) | I am a moderate conservative and there are some things about anti-gay rights that I don't support. | Everything. |

# Who scores higher on political knowledge?

Both respondents scored equally on the conventional political knowledge measure in the 2012 ANES!

# Who scores higher on political knowledge?

| | **Respondent A** | **Respondent B** |
|---|---|---|
| Obama (+) | I think he is honest, has good intentions. | He's got a hot wife. Sound mind. Bill Clinton, the best president we ever had, is voting for him. He's not in any type of scandal and is the perfect president for us now. |
| Obama (-) | I don't feel he is up for the job, he doesn't really know how to get things accomplished from idea to actual reality. | |
| Romney (+) | He comes across as an honest person and I feel that financially he would be better for the country. | |
| Romney (-) | I am a moderate conservative and there are some things about anti-gay rights that I don't support. | Everything. |

# Measuring text-based sophistication

1. Relative length $= \dfrac{\log\left(\sum_{j=1}^{J} n_{ij}\right)}{\max\left[\log\left(\sum_{j=1}^{J} n_{ij}\right)\right]}$

---

| | |
|---|---|
| $n$ | Word count |
| $i$ | Individual respondent |
| $j \in \{1, ..., J\}$ | Open-ended items |

# Measuring text-based sophistication

1. Relative length $= \dfrac{\log\left(\sum_{j=1}^{J} n_{ij}\right)}{\max\left[\log\left(\sum_{j=1}^{J} n_{ij}\right)\right]}$
2. Topic diversity $= 1 - \dfrac{\sum_{k_1=1}^{K}\sum_{k_2=1}^{K} |\theta_{ik_1} - \theta_{ik_2}|}{2K\sum_{k_1=1}^{K}\theta_{ik_1}}$

---

| | |
|---|---|
| $n$ | Word count |
| $i$ | Individual respondent |
| $j \in \{1, ..., J\}$ | Open-ended items |
| $k \in \{1, ..., K\}$ | Topics (estimated via STM, Roberts et al. 2014) |
| $\theta_{ik}$ | Predicted proportion of topic $k$ in the collection of responses by individual $i$ |

# Measuring text-based sophistication

1. Relative length $= \dfrac{\log\left(\sum_{j=1}^{J} n_{ij}\right)}{\max\left[\log\left(\sum_{j=1}^{J} n_{ij}\right)\right]}$
2. Topic diversity $= 1 - \dfrac{\sum_{k_1=1}^{K}\sum_{k_2=1}^{K} \left|\theta_{ik_1} - \theta_{ik_2}\right|}{2K\sum_{k_1=1}^{K}\theta_{ik_1}}$
3. Opinionation $= 1 - \dfrac{\sum_{j_1=1}^{J}\sum_{j_2=1}^{J} \left|p_{ij_1} - p_{ij_2}\right|}{2J\sum_{j_1=1}^{J} p_{ij_1}}$

---

| | |
|---|---|
| $n$ | Word count |
| $i$ | Individual respondent |
| $j \in \{1, ..., J\}$ | Open-ended items |
| $k \in \{1, ..., K\}$ | Topics (estimated via STM, Roberts et al. 2014) |
| $\theta_{ik}$ | Predicted proportion of topic $k$ in the collection of responses by individual $i$ |
| $p_{jk}$ | Proportion of words in the response of individual $i$ to question $j$ relative to the overall size of the individual's response |

# Data

- Main analysis: 2012 American National Election Study (ANES)
    - $N = 5914$ ($2054$ f2f $+$ $3860$ online)
    - Non-response: 417, Spanish: 228

# Data

- Main analysis: 2012 American National Election Study (ANES)
    - N = 5914 (2054 f2f + 3860 online)
    - Non-response: 417, Spanish: 228
- Additional analyses and replication (in paper):

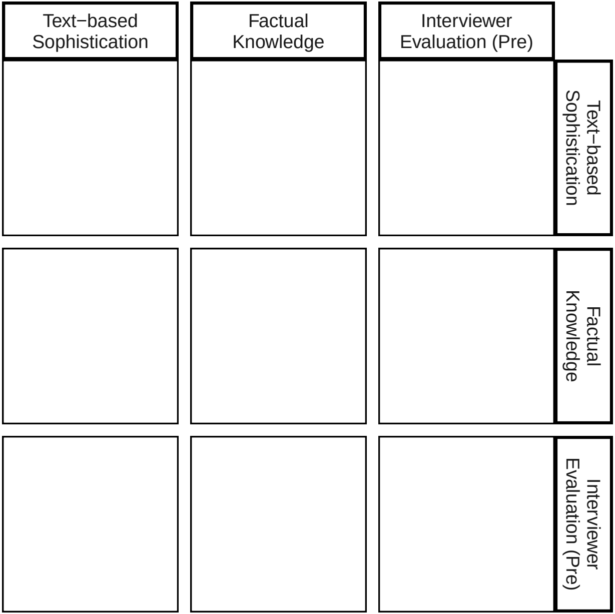
# Data

- Main analysis: 2012 American National Election Study (ANES)
  - N = 5914 (2054 f2f + 3860 online)
  - Non-response: 417, Spanish: 228
- Additional analyses and replication (in paper):
  - 2015 YouGov Survey:
    N = 1000 (online) - 48 (non-response)

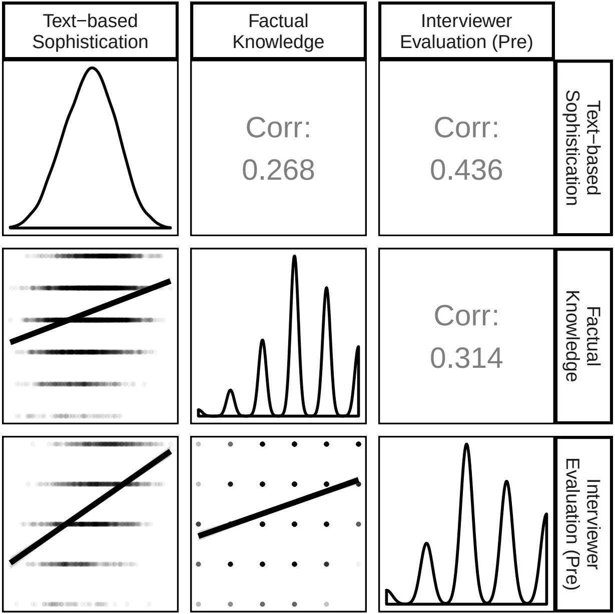
# Data

- Main analysis: 2012 American National Election Study (ANES)
    - N = 5914 (2054 f2f + 3860 online)
    - Non-response: 417, Spanish: 228
- Additional analyses and replication (in paper):
    - 2015 YouGov Survey:
      N = 1000 (online) - 48 (non-response)
    - 2008–2012 Swiss Referendum Survey (Colombo, 2016):
      N = 26,621 (phone) - 4,917 (non-response)

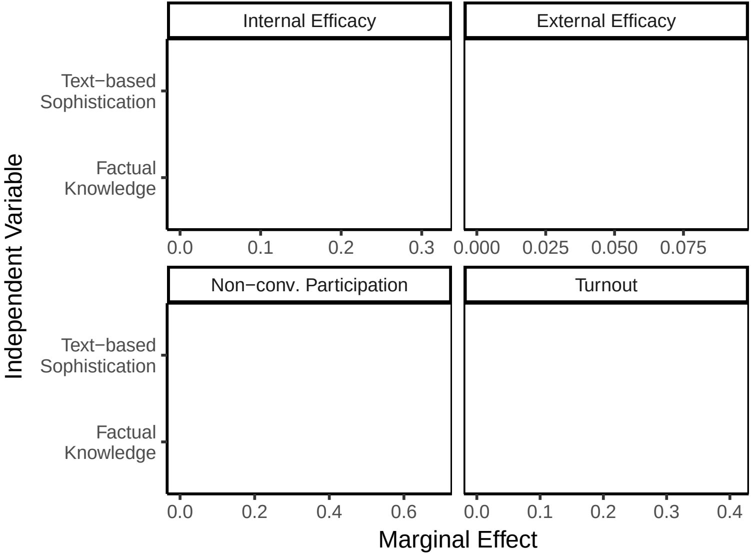
# Comparison with conventional measures (ANES 2012)



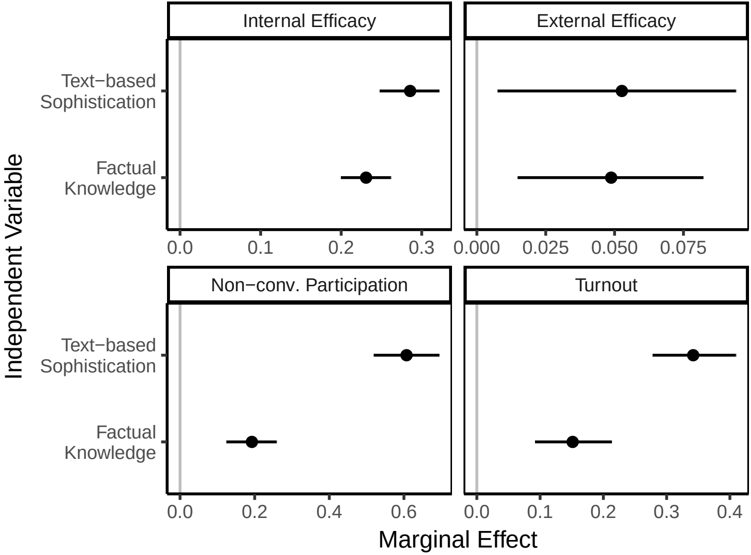
# Comparison with conventional measures (ANES 2012)



# Validation: Engagement and participation (ANES 2012)



# Validation: Engagement and participation (ANES 2012)



# Other validation results

- Text-based sophistication is associated with ...

# Other validation results

- Text-based sophistication is associated with ...
  ... more precise candidate and party placements on multiple policy issues.

# Other validation results

- Text-based sophistication is associated with ...
  - ... more precise candidate and party placements on multiple policy issues.
  - ... higher likelihood that citizens voted according to their initial intention at the time of the pre-election interview.

# Other validation results

- Text-based sophistication is associated with **...**
  - **...** more precise candidate and party placements on multiple policy issues.
  - **...** higher likelihood that citizens voted according to their initial intention at the time of the pre-election interview.
  - **...** more accurate information retrieval from news articles.

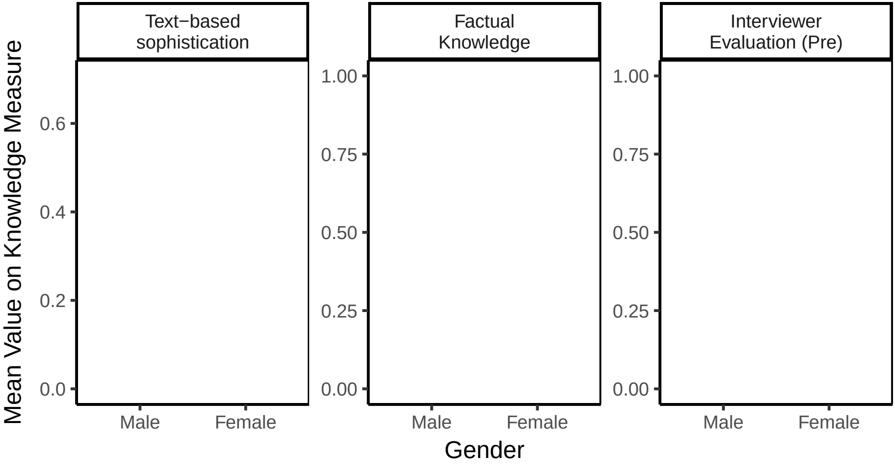
# Other validation results

- Text-based sophistication is associated with **...**
  - **...** more precise candidate and party placements on multiple policy issues.
  - **...** higher likelihood that citizens voted according to their initial intention at the time of the pre-election interview.
  - **...** more accurate information retrieval from news articles.
  - **...** manual coding of individual levels of justification (Colombo, 2016).

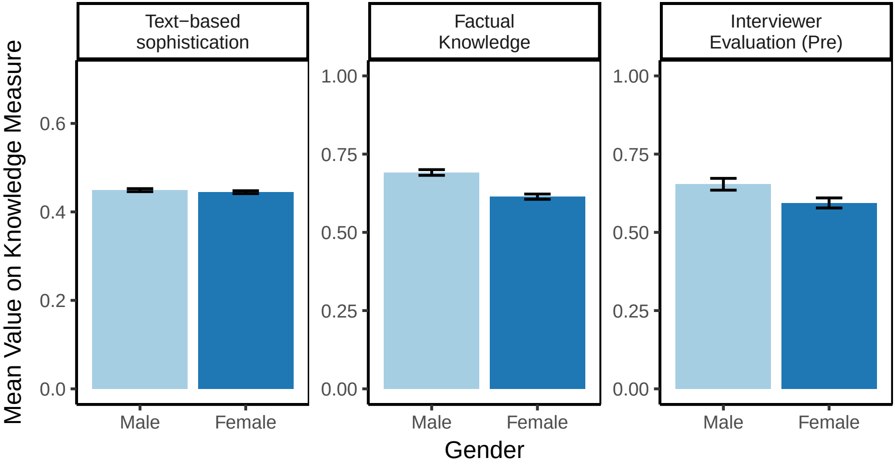
# Other validation results

- Text-based sophistication is associated with ...
  - ... more precise candidate and party placements on multiple policy issues.
  - ... higher likelihood that citizens voted according to their initial intention at the time of the pre-election interview.
  - ... more accurate information retrieval from news articles.
  - ... manual coding of individual levels of justification (Colombo, 2016).
- But: conventional measures of factual knowledge are, too...

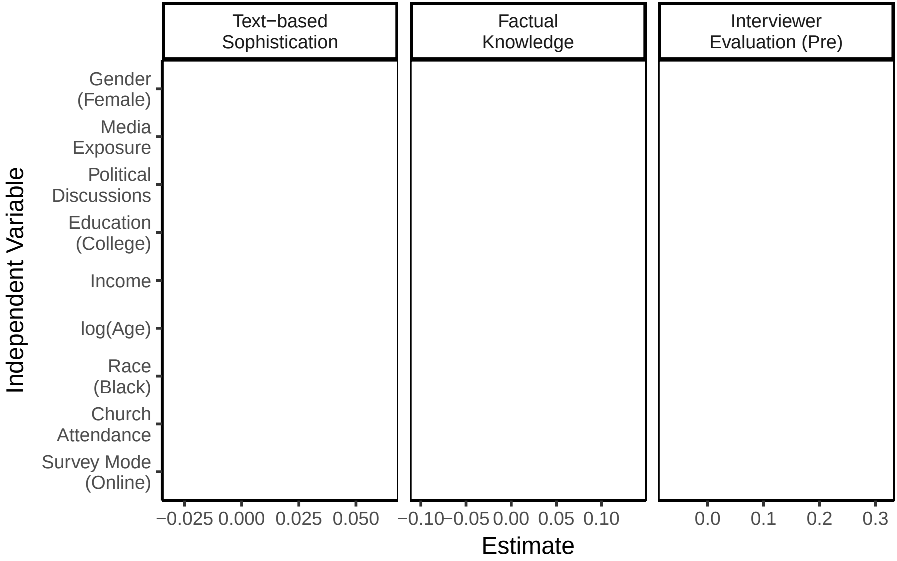
# Application: Assessing the gender gap I



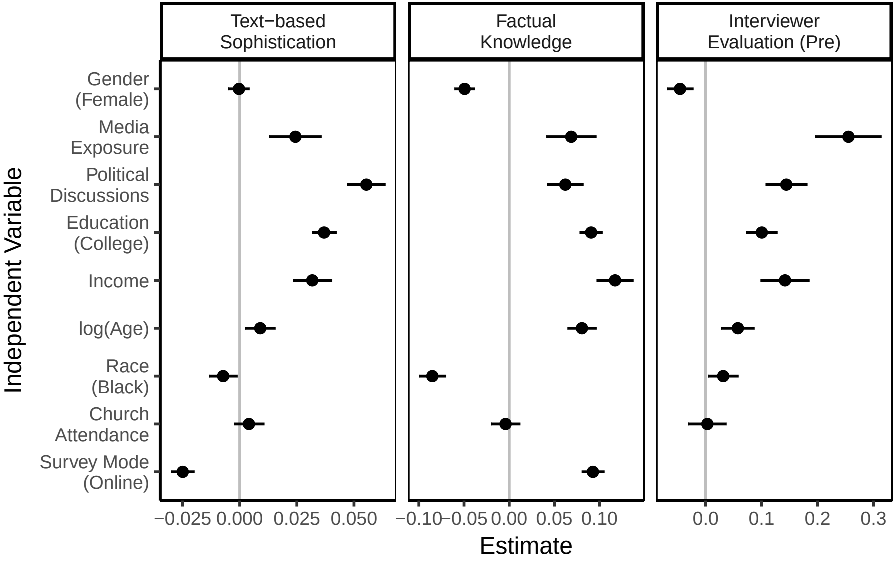
# Application: Assessing the gender gap I



# Application: Assessing the gender gap II



# Application: Assessing the gender gap II



# Conclusion

- We observe theoretically meaningful variation in the complexity of verbatim open-ended responses.

# Conclusion

- We observe theoretically meaningful variation in the complexity of verbatim open-ended responses.
- By directly examining how individuals justify their attitudes, we can measure sophistication related to specific political tasks.

# Conclusion

- We observe theoretically meaningful variation in the complexity of verbatim open-ended responses.
- By directly examining how individuals justify their attitudes, we can measure sophistication related to specific political tasks.
- Text-based sophistication is conceptually closer to the structure of belief systems than conventional measures (e.g., Tetlock, 1983; Luskin, 1987).

# Conclusion

- We observe theoretically meaningful variation in the complexity of verbatim open-ended responses.
- By directly examining how individuals justify their attitudes, we can measure sophistication related to specific political tasks.
- Text-based sophistication is conceptually closer to the structure of belief systems than conventional measures (e.g., Tetlock, 1983; Luskin, 1987).

- Women might score lower than men on factual knowledge about political institutions, but there are no differences in the sophistication of expressed political attitudes.

# Conclusion

- We observe theoretically meaningful variation in the complexity of verbatim open-ended responses.
- By directly examining how individuals justify their attitudes, we can measure sophistication related to specific political tasks.
- Text-based sophistication is conceptually closer to the structure of belief systems than conventional measures (e.g., Tetlock, 1983; Luskin, 1987).

- Women might score lower than men on factual knowledge about political institutions, but there are no differences in the sophistication of expressed political attitudes.

`#WomenAlsoKnowStuff`

# Who scores higher on political knowledge?

| | **Respondent A** | **Respondent B** |
|---|---|---|
| Obama (+) | I think he is honest, has good intentions. | He's got a hot wife. Sound mind. Bill Clinton, the best president we ever had, is voting for him. He's not in any type of scandal and is the perfect president for us now. |
| Obama (-) | I don't feel he is up for the job, he doesn't really know how to get things accomplished from idea to actual reality. | |
| Romney (+) | He comes across as an honest person and I feel that financially he would be better for the country. | |
| Romney (-) | I am a moderate conservative and there are some things about anti-gay rights that I don't support. | Everything. |

Thank you very much for your attention!

Manuscript and code available at:
https://github.com/pwkraft/knowledge

Comments, questions?
patrick.kraft@stonybrook.edu

# References

Colombo, Céline. 2016. "Justifications and Citizen Competence in Direct Democracy: A Multilevel Analysis." *British Journal of Political Science*: 1–20.

Luskin, Robert C. 1987. "Measuring political sophistication." *American Journal of Political Science* 31 (4): 856–899.

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.