Is Mathilda Playing it Safe? Gender in Computational Text Analysis Methods *

Alona Dolinsky *University College Dublin*
Mariken A.C.G. van der Velden *Vrije Universiteit Amsterdam*

The editors of top-journals in political science have investigated the role of the review process finding no indication for gender bias in the editorial process. They suggest that the submission pools are distorted by gender. We test some pathways that could distort the submission pool: a) playing it safe due to the gender perception gap, and b) the Mathilda effect. Papers using Computational Text Analysis Methods are more likely to be published in journals with a 'masculinized' perception gap. When women are aiming for these journals, they might 'play it safe' by conducting more validation checks than their male colleagues. Moreover, embracing the Mathilda effect – i.e. the systematic under-recognition of female scientists – women scholars are more likely to indicate that a) there are important training needs in more areas; and b) they themselves need (further) training in computational methods and use these reasons not to publish employing these methods. We test these claims using a) an unique content analysis of research articles published in the top 20 journals in communication science, political science, sociology and psychology between 2016 and 2020, identifying all 854 articles that involved some form of quantitative textual analysis; and b) an expert survey with all authors of quantitative text analytic research identified via said content analysis, which inquired about researchers' considerations and concerns in the application of computational text analytic strategies.

Keywords: Experiment, Populist Attitudes

Introduction

Motivation: why do we care

Gatto et al. (2020) demonstrates that while women cover significantly fewer methods courses in their doctoral training, when they do participate in methods training, they, however, show similar levels of method employment to their male colleagues. Yet, Teele and Thelen (2017) demonstrate that women are less likely to publish studies using quantitative and computational methods (see Maliniak, Powers, and Walter (2013) for a similar argument in International Relations). Brown et al. (2020) demonstrate that the *gender submission gap* is accompanied by a *gender perception gap*: Women report that they are more likely to submit to and get published in some journals, whereas men report as such with regard to other journals. Importantly, these gaps are observed even among scholars with the same methodological (i.e., quantitative or qualitative) approach. The editors of top-journals in politicial science have investigated the role of the review process finding no indication for gender bias in the editorial process (Breuning et al. 2018; König and Ropers 2018; Brown and Samuels 2018). The authors all suggest that the submission pools are distorted by gender.

We test X pathways that could distort the submission pool: a) playing it safe due to the gender perception gap, and b) the Mathilda effect. Papers using CTAM are more likely to be published

*Replication files are available on the author's Github account (https://anonymous.4open.science/r/gender-ctam-CBA3/README.md). **Current version**: February 24, 2022; **Corresponding author**: Mariken A.C.G. van der Velden

in journals with a *"masculinized" perception gap*. When women are aiming for these journals, it could be that they "play it safe' ' by conducting more validation checks than their men. Moreover, embracing of the *Mathilda effect* (Rossiter 1993) – i.e. the systematic underrecognition of female scientists – women scholars are more likely to indicate that a) there are important training needs in more areas; and b) they themselves need (further) training in computational methods and use these reasons not to publish employing these methods.

Play it Safe hypothesis (*H1a*): Women scholars are more likely to play it safe and indicate more validation strategies than men scholars.

Play it Safe hypothesis (*H1b*): Women scholars are more likely to play it safe and indicate more challenges as reasons **not** to use CTAM than men scholars.

Mathilda effect hypothesis (*H2a*): Women scholars are more likely to indicate a higher number of important training needs than men scholars.

Mathilda effect hypothesis (*H2b*): Women scholars are more likely to indicate that they themselves require (further advanced) training than men scholars.

Data & Results

We have condected an expert survey, inviting all scholars published a scientific article between January 2016 and September 2020 (`N = 45,437`). Using a keyword search on the Web of Science, we then identified a total of 7,296 *potentially* relevant articles whose abstracts referred to some kind of textual contents or text analytic procedures. We then accessed the full text of these articles to determine whether the presented research included any form of quantitative textual analysis. Quantitative textual analysis was defined broadly to include any form of processing natural language that identified specific kinds of textual contents with the purpose of classification and quantitative analysis. Analyses that relied solely on metadata or pre-existing classifications were excluded, as were investigations accessing only formal properties of the sampled texts (e.g., length). We included analyses of multi-modal media (e.g., posters, television) as long as textual contents were informative toward classification. Purely methodological contributions discussing specific potentials or limitations of available methods were excluded, unless they included applied demonstrations wherein actual textual data was processed. Articles were considered relevant as soon as they used any form of quantitative textual analysis, even if it was used merely in an auxiliary capacity (e.g., a content analysis to identify frames to be used in an experiment; sentiment analyses of open-ended survey responses). This screening yielded a total of `N = 854` articles, for which the authors were looked up. This yielded us with 1,653 identifiable and working email-addresses.

The experts have been invited to the questionnaire on March 4th of 2021, and received two reminders, each approximately a week after our last message (respectively on March 11th and March 16th of 2021). This yielded a responses of 433 responses (i.e. response rate of `25%`). The study has been approved by the Research Ethics Review Committee of the *Vrije Universiteit Amsterdam*, *Hebrew University Jerusalem Isreal*, and *University College Dublin*.

Figure 1: Variables under Study

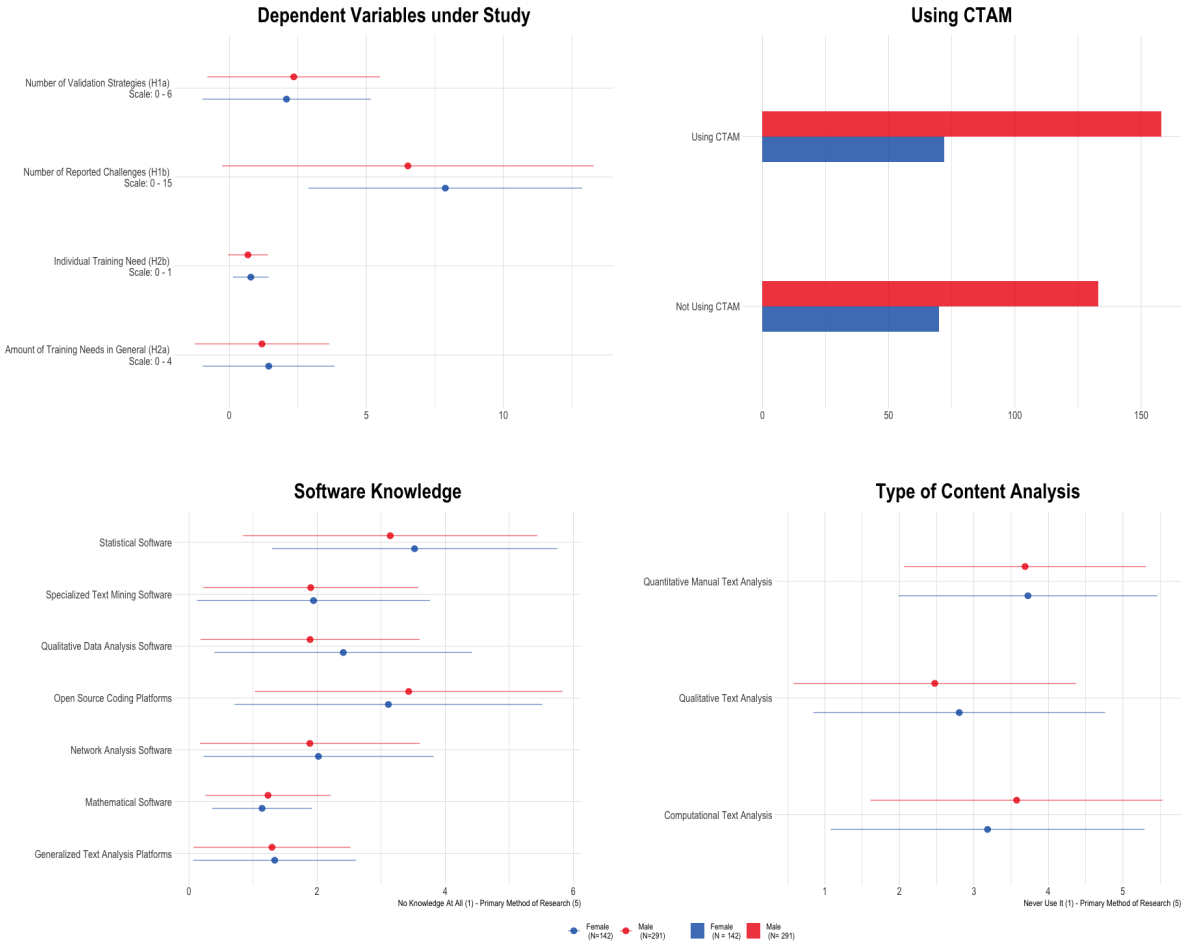


Figure 2: Results Pre-Registered Hypotheses

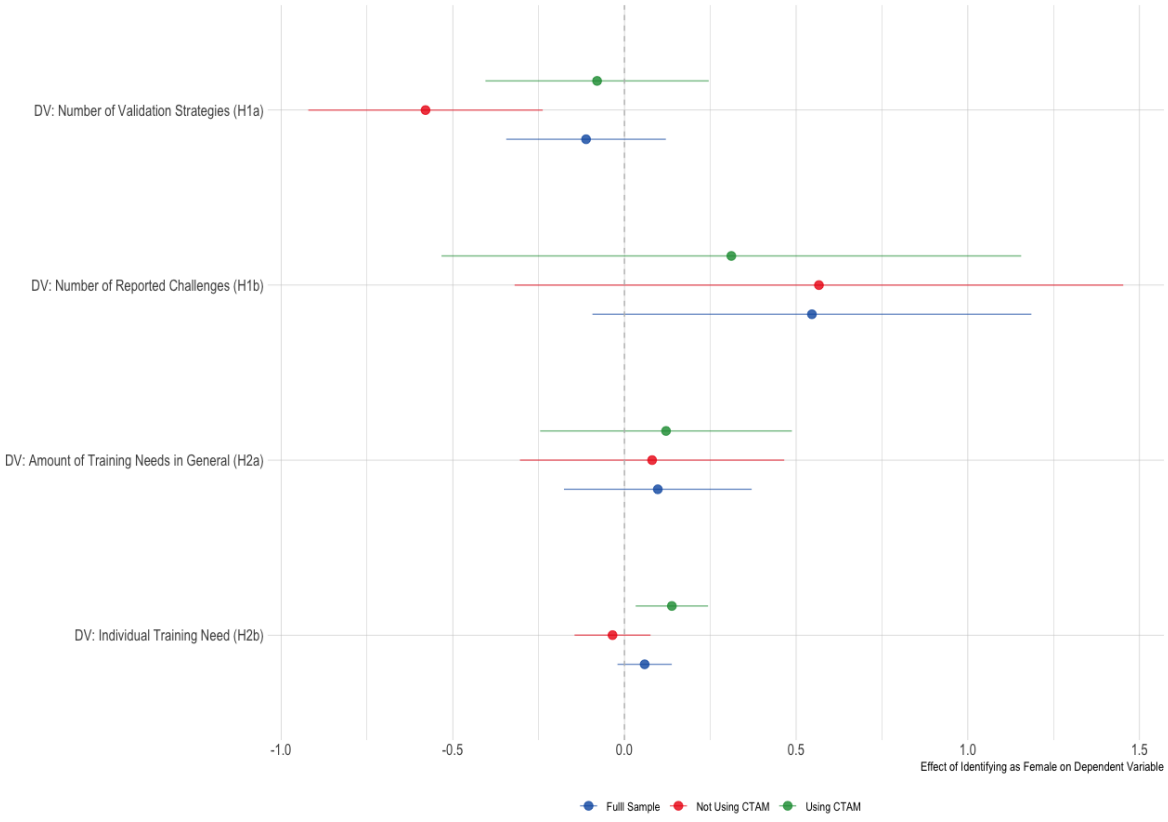


Figure 3: Results Explorative Analyses

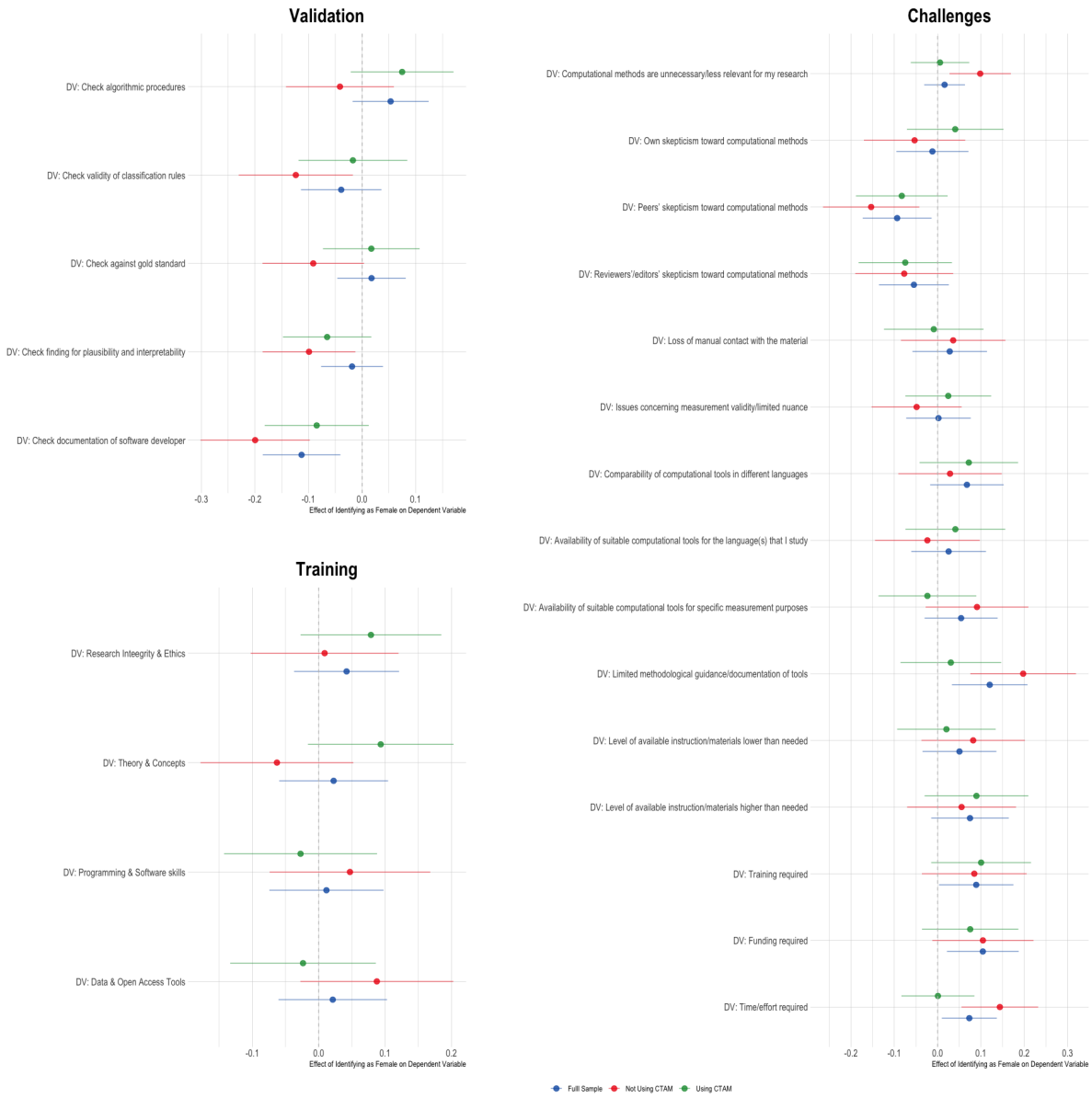
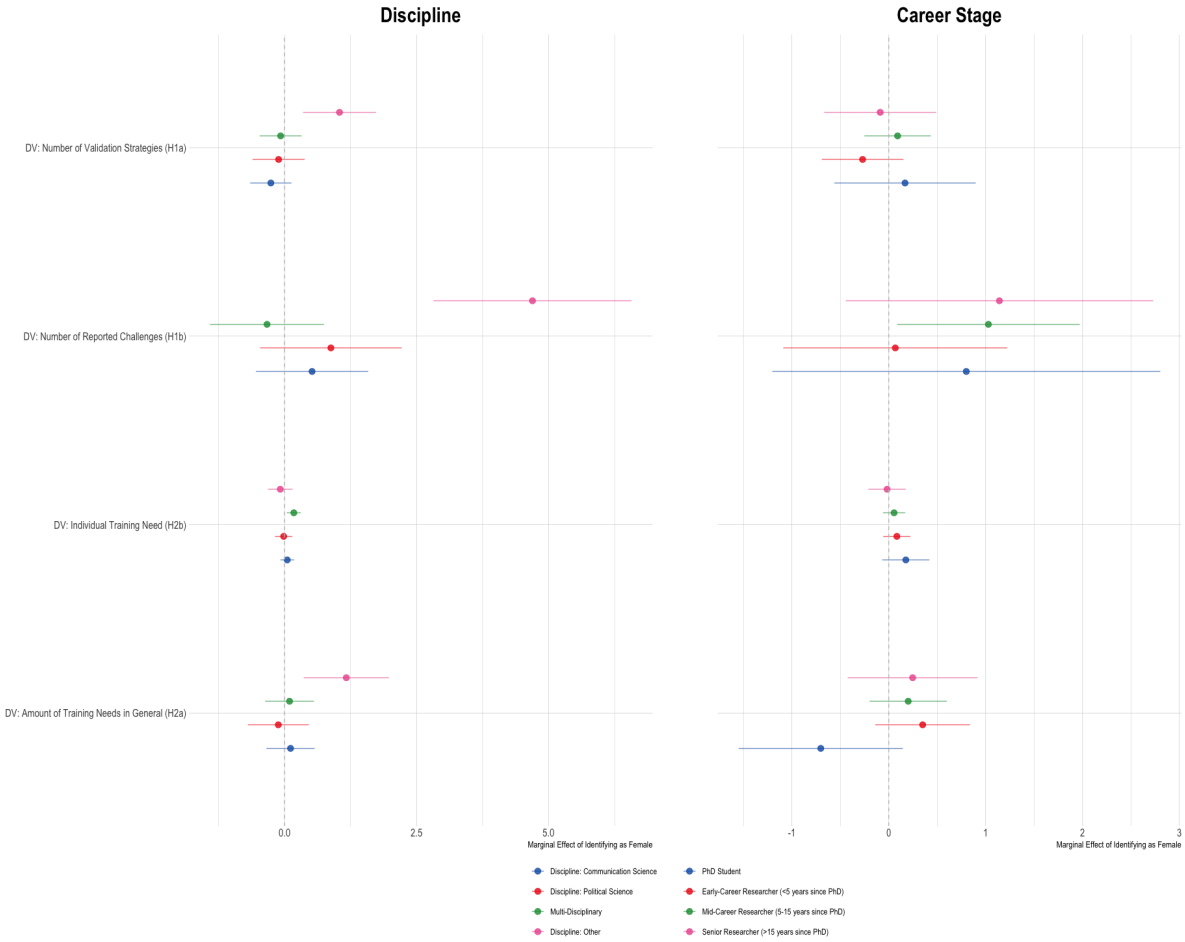


Figure 4: Results Explorative Analyses



Conclusion

References

Breuning, Marijke, Benjamin Isaak Gross, Ayal Feinberg, Melissa Martinez, Ramesh Sharma, and John Ishiyama. 2018. "Clearing the Pipeline? Gender and the Review Process at the American Political Science Review." *PS: Political Science& Politics* 51 (3): 629–34. https://doi.org/10.1017/S1049096518000069.

Brown, Nadia E., Yusaku Horiuchi, Mala Htun, and David Samuels. 2020. "Gender Gaps in Perceptions of Political Science Journals." *PS: Political Science &Amp; Politics* 53 (1): 114–21. https://doi.org/10.1017/S1049096519001227.

Brown, Nadia E., and David Samuels. 2018. "Beyond the Gender Citation Gap: Comments on Dion, Sumner, and Mitchell." *Political Analysis* 26 (3): 328–30. https://doi.org/10.1017/pan.2018.14.

Gatto, Malu AC, Anita R Gohdes, Denise Traber, and Mariken ACG Van Der Velden. 2020. "Selecting in or Selecting Out? Gender Gaps and Political Methodology in Europe." *PS: Political Science & Politics* 53 (1): 122–27. https://doi.org/10.1017/S1049096519001288.

König, Thomas, and Guido Ropers. 2018. "Gender and Editorial Outcomes at the American Political Science Review." *PS: Political Science &Amp; Politics* 51 (4): 849–53. https://doi.org/10.1017/S1049096518000604.

Maliniak, Daniel, Ryan Powers, and Barbara F. Walter. 2013. "The Gender Citation Gap in International Relations." *International Organization* 67 (4): 889–922. https://doi.org/10.1017/S0020818313000209.

Rossiter, Margaret W. 1993. "The Matthew Matilda Effect in Science." *Social Studies of Science* 23 (2): 325–41. https://doi.org/10.1177/030631293023002004.

Teele, Dawn Langan, and Kathleen Thelen. 2017. "Gender in the Journals: Publication Patterns in Political Science." *PS: Political Science &Amp; Politics* 50 (2): 433–47. https://doi.org/10.1017/S1049096516002985.