Pre-Analysis Plan

16 April 2021

Contents

1 Hypotheses

2 Research Desing and Protocol

2.1 Sample

I will conduct this survey experiment in the Netherlands in April 2021. Our sample, recruited through KiesKompas, will consist of 2,000 participants (based on the power analysis presented below) of 18 years and older. The study has been approved by the Research Ethics Review Committee of the *Vrije Universiteit Amsterdam* (see the approval here. To ensure good quality of our data, two attention checks (discussed in more detail below) are included. Each respondent failing the attention check will be excluded and replaced with another 'good' response.

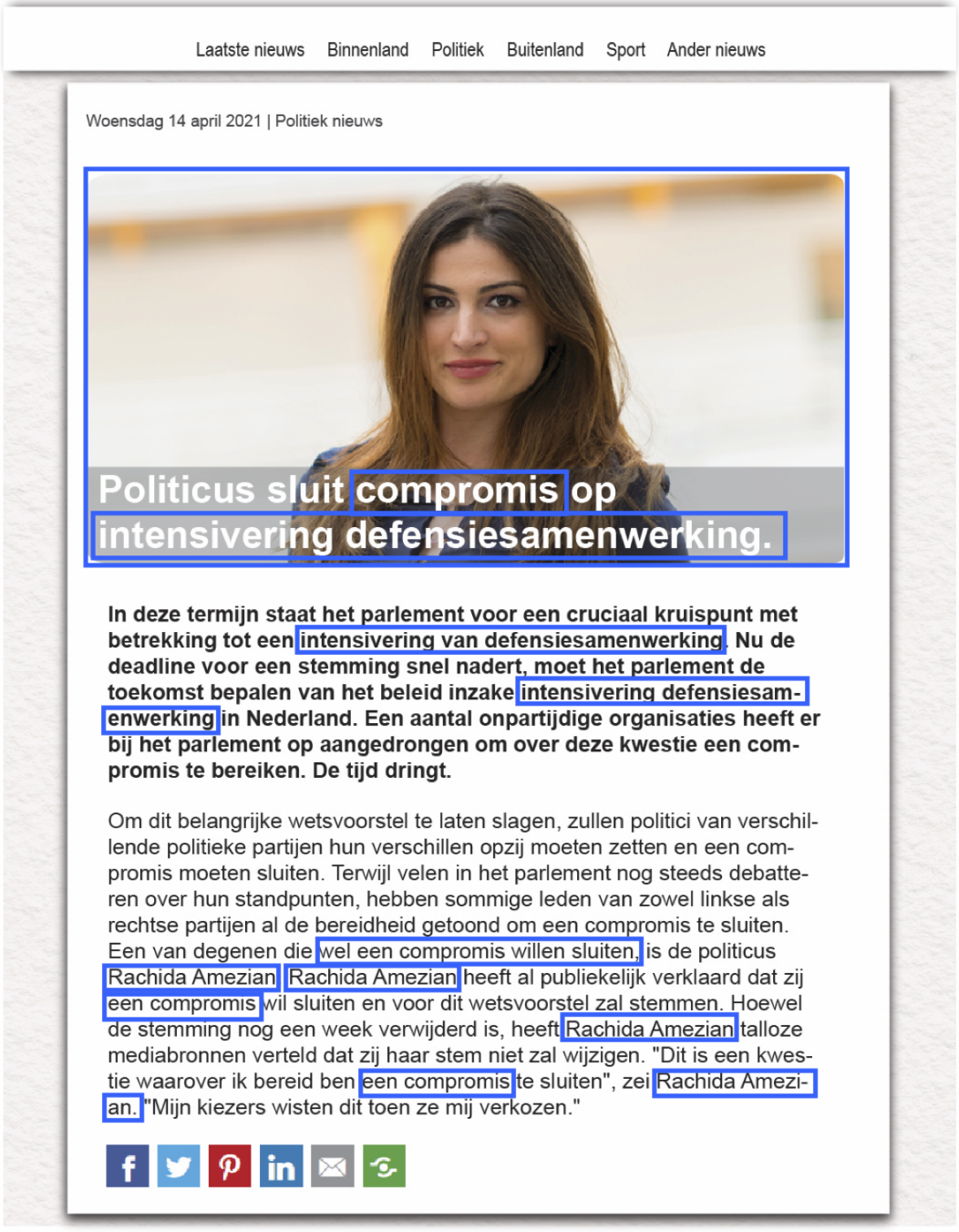
2.2 Experimental Protocol

The study is conducted online and in Dutch. Participants are told that they are taking part in a survey to get an overview of how Dutch people form their views on politics. After reading an informed consent message participants are forwarded to the main questionnaire (or the survey will be terminated if they do not agree to the consent form).

First, participants complete a set of background variables on their stances on the political issues used in the experiment (i.e. corona-measures, defense, education, immigration) and on their attitude towards women in politics – the full codeboek can be viewed here. The pre-treatment block ends with one of the two attention checks included in this survey. When participants fail this attention check, a warning appears asking them to read the question again carefully and to answer again. Only when they have answered it correctly, they enter the first round of the experiment. After each round of the experiment, some filler questions are asked about the respondents demographics. The stimuli in the experiment are news messages in the same style as the Dutch news website nu.nl. In these news messages, we manipulate: a) the *gender* of the politician (male vs. female); b) the *migration background* of the politician (based on a Arabic sounding name (Rachid(a) Amezian) or a native Dutch sounding name (Karel/Karin van der Kleijn)); and c) whether the politician *striked a compromise or not*. This creates a full `2*2*2` factorial experiment with four rounds. Every round, the news message covers a new issue. The issues are: "defense" (increase cooperation of the defense units), "immigration" (immigration stop in western countries), "education" (expansion creative elementary school activities), "COVID-19" (emergency legislation to combat COVID-19). For every round, the politicians name, and thereby gender and immigration background, are manipulated in the text and in the picture as well as the decision to strike a compromise. For an illustration of the stimulus material, see Figure 1. All manipulated elements are within the blue boxes.

The text of the treatment is a Dutch translation of the stimulus material developed by Bauer, Yong, and Krupnikov (2017). The authors have pre-tested the text of the stimuli material several times (see the Web Appendix of Bauer, Yong, and Krupnikov 2017). The photo's of their stimuli material can unfortunately not be used, because a) Bauer, Yong, and Krupnikov (2017) only investigate the difference between men and women politicians, using solely photo"s of caucasians; and b) the American flag was present in their material. Instead, I have used photo's from local politicians of the city of The Hague that are not know to the general public. The men politicians are in real life representatives of the liberal-progressive party D66 (man with immigrant background) and of the christian democratic party CDA (man without immigrant background). The women politicians are in real life representatives of a right-wing populist party *Hart voor Den Haag/ Groep de Mos* (woman with migration background) and of the social democratic party PvdA (woman without immigrant background). Thereby we have a spread over the ideological different parties present in Dutch politics. The names for the politicians without immigration background are *Karel* (man) and *Karin* (woman) *van der Kleijn*, which are typical native Dutch names. For the politicians with an immigration background, we picked *Rachid* (man) and *Rachida* (woman) *Amezian*.

Figure 1: Annotated Example of Stimulus Material



2.3 Power Analysis

As detailed in Section 4, I conduct an OLS regression within a multiverse approach (Simonsohn, Simmons, and Nelson 2019; Steegen et al. 2016) using the four post-treatment questions on the politician's a) leadership evaluaton (scale); b) general favorability; c) representative quality; and d) career perspective as dependent variable and the three manipulations (gender, migration background and whether or not the politician is willing to strike a compromise) as independent variables. Each hypothesis is tested separately for all four issues. To calculate power for the hypotheses, the R package `DeclareDesign` is used (Blair et al. 2019). Based on the study of Bauer, Yong, and Krupnikov (2017), the effect sizes are between `0.2` and `0.1` – i.e. a small effect visualized by the purple lines in Figure 2. The hypothesis are directional, Figure 2 therefore displays one-tailed tests with α `= 0.05`. The power analysis shows that testing hypotheses 1 and 2 requires a sample size of `1,800` participants (x-axis) to reach `95%` power (black dashed line, in the left-panel of Figure 2). Note that if the effect size is bigger then `0.2` (i.e. `0.3`, `0.4`, `0.5`, or `0.6`) smaller sample sizes are sufficient to reach `95\%` power. To test H3, the combination of gender and migration background, we can detect a small effect `b = 0.2` (purple line) with a power of `80%` (gray dashed line) and a one-tailed test of signficance at α `=0.05` with a sample of `1,800` participants (Right-Panel of Figure 2). At `2,000` participants, with a small effect and one-tailed test of significance at α `=0.05`, I reach power of `88%`. A sample of `2,000` participants therefore gives me suficient power to test Hypotheses 1 and 2, but a probability of `12%` for a Type II error remains when testing Hypotheses 3. I will therefore test all hypotheses first by issue and second, by pooling our data across issues. As Figure 3 demonstrates, this will give sufficient power even if we do two-tailed testing.

```
source(here::here("docs/pap/poweranalysis.R"))
```

Figure 2: Power Analysis

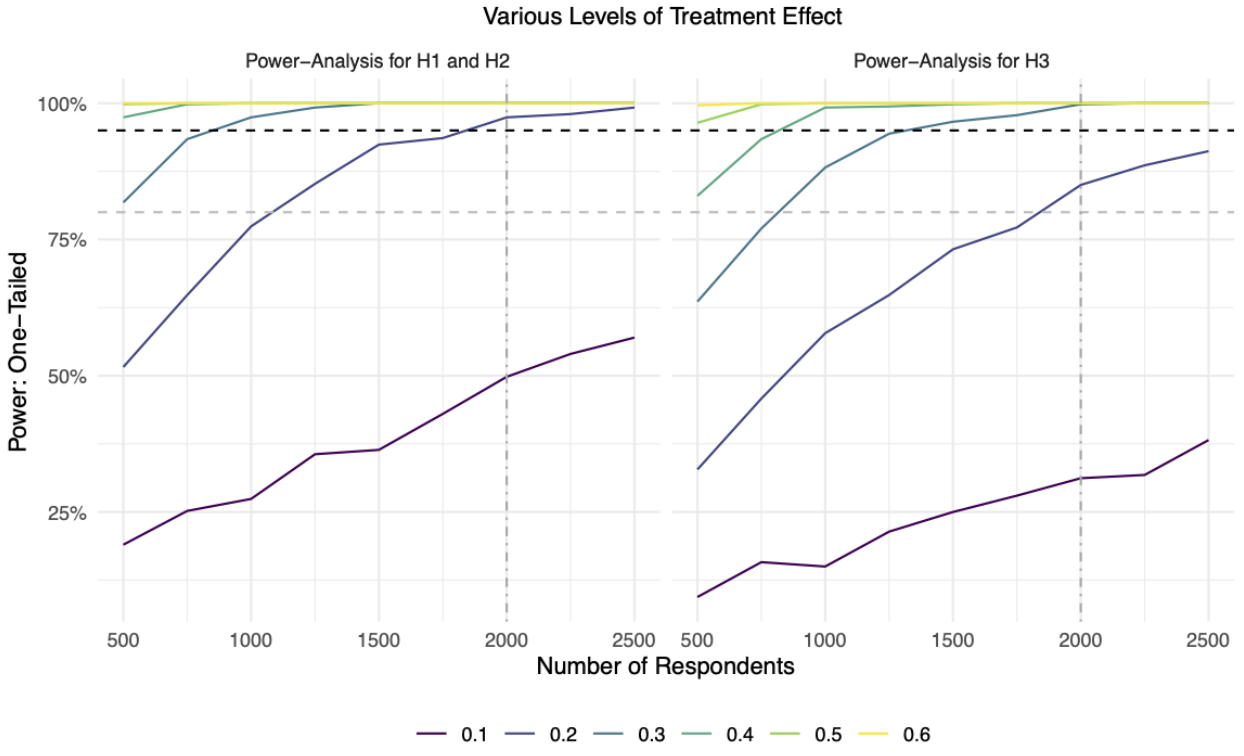
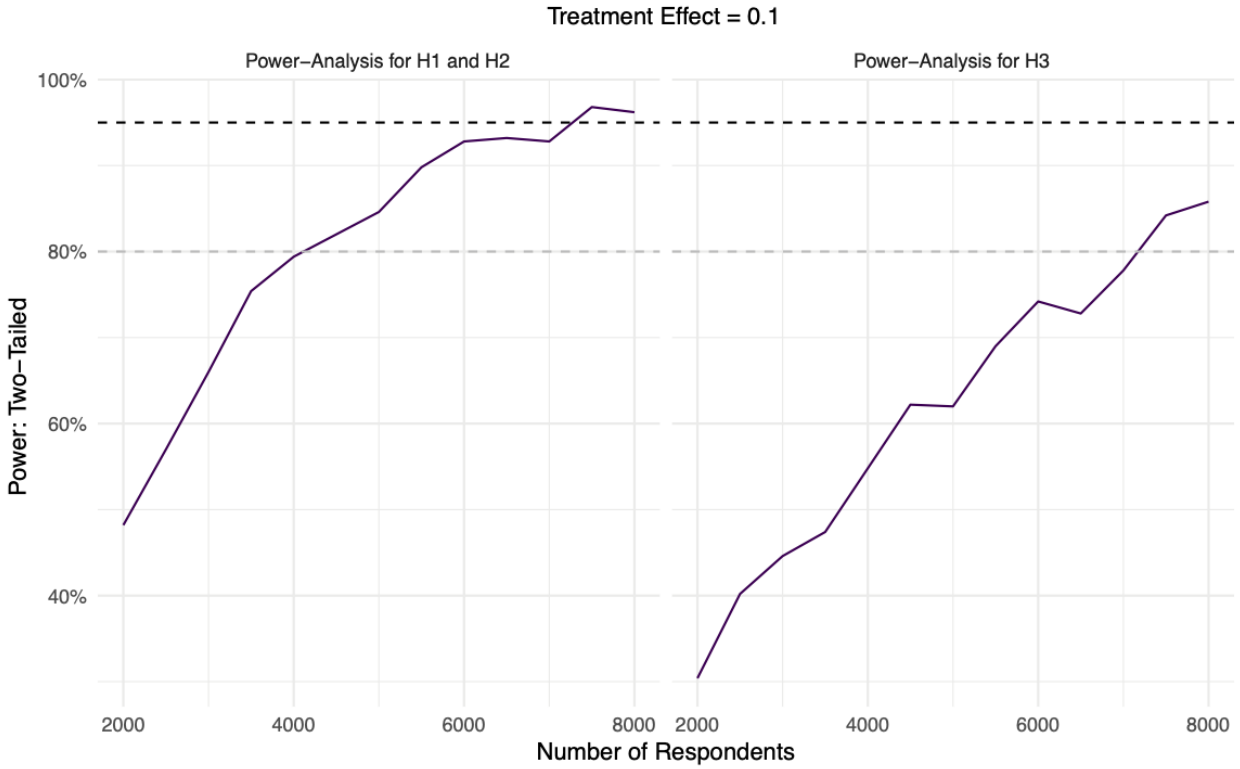


Figure 3: Power Analysis - Pooled



3 Measures

3.1 Dependent Variables

3.2 Control Variables

3.3 Attention Checks

3.4 Exclusion Criteria

4 Analysis

4.1 Hypothesis 1

4.2 Hypothesis 2

4.3 Hypothesis 3

5 Stimulus Material

References

Bauer, Nichole M, Laurel Harbridge Yong, and Yanna Krupnikov. 2017. "Who Is Punished? Conditions Affecting Voter Evaluations of Legislators Who Do Not Compromise." *Political Behavior* 39 (2). Springer: 279–300. https://doi.org/10.1007/s11109-016-9356-6.

Blair, Graeme, Jasper Cooper, Alexander Coppock, and Macartan Humphreys. 2019. "Declaring and Diagnosing Research Designs." *American Political Science Review* 113 (3). Cambridge University Press: 838–59. https://doi.org/10.1017/S0003055419000194.

Simonsohn, Uri, Joseph P Simmons, and Leif D Nelson. 2019. "Specification Curve: Descriptive and Inferential Statistics on All Reasonable Specifications." *Available at SSRN 2694998*. https://doi.org/10.2139/ssrn.2694998.

Steegen, Sara, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. 2016. "Increasing Transparency Through a Multiverse Analysis." *Perspectives on Psychological Science* 11 (5). Sage Publications Sage CA: Los Angeles, CA: 702–12. https://doi.org/10.1177/1745691616658637.