Pre-Analysis Plan: Ideological Biases of the Crowd?

02 June, 2022 - 21:53:32 (CEST)

Contents

1 Expectations

Ideological Bias hypothesis (*H1a*): The larger the ideological distance between respondent and the party, the less likely respondents code stances according to the party's position.

Ideological Bias hypothesis (*H1b*): The effect of H1a is stronger for the underspecified condition.

Ideological Overinterpretation hypothesis (*H2a*): The larger the ideological distance between respondent and the party, the more likely respondents interpret underspecified sentences as stance.

Political Knowledge Overinterpretation hypothesis (*H2b*): The more political knowledge, the more likely people interpret underspecified sentences as stance.

Masking Solution hypothesis (*H3a*): Masking reduces the effect of respondents' ideological position for coding stances according to the party's position.

Masking Solution hypothesis (*H3b*): Masking reduces the effect of respondents' level of political knowledge for coding stances according to the party's position.

2 Research Desing and Protocol

2.1 Sample

We will conduct this survey experiment in the Netherlands in April 2022. The sample, recruited through KiesKompas, will consist of 2,000 participants (based on the power analysis presented in Figure 1) of 18 years and older. Kieskompas works with non-random opt-in respondents. Therefore, I measure many demographic background variables (see Section 3.2). Balance checks will be conducted to demonstrate whether certain categories are over represented in a certain experimental group. 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, one attention check (discussed in more detail in Section 3.3) is included.

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 pre-treatment variables (i.e. vote recall, issue positions of the issues used in the experiment (Climate, EU, Immigration, and Tax), ideology, and political knowledge). This block ends with the attention check 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. Therefter, participants see the experimental stimuli. The stimuli in the experiment are to show respondents fully specified or underspecified sentences on the issue position of a political actor or a masked political actor (see Section Treatment for more details).

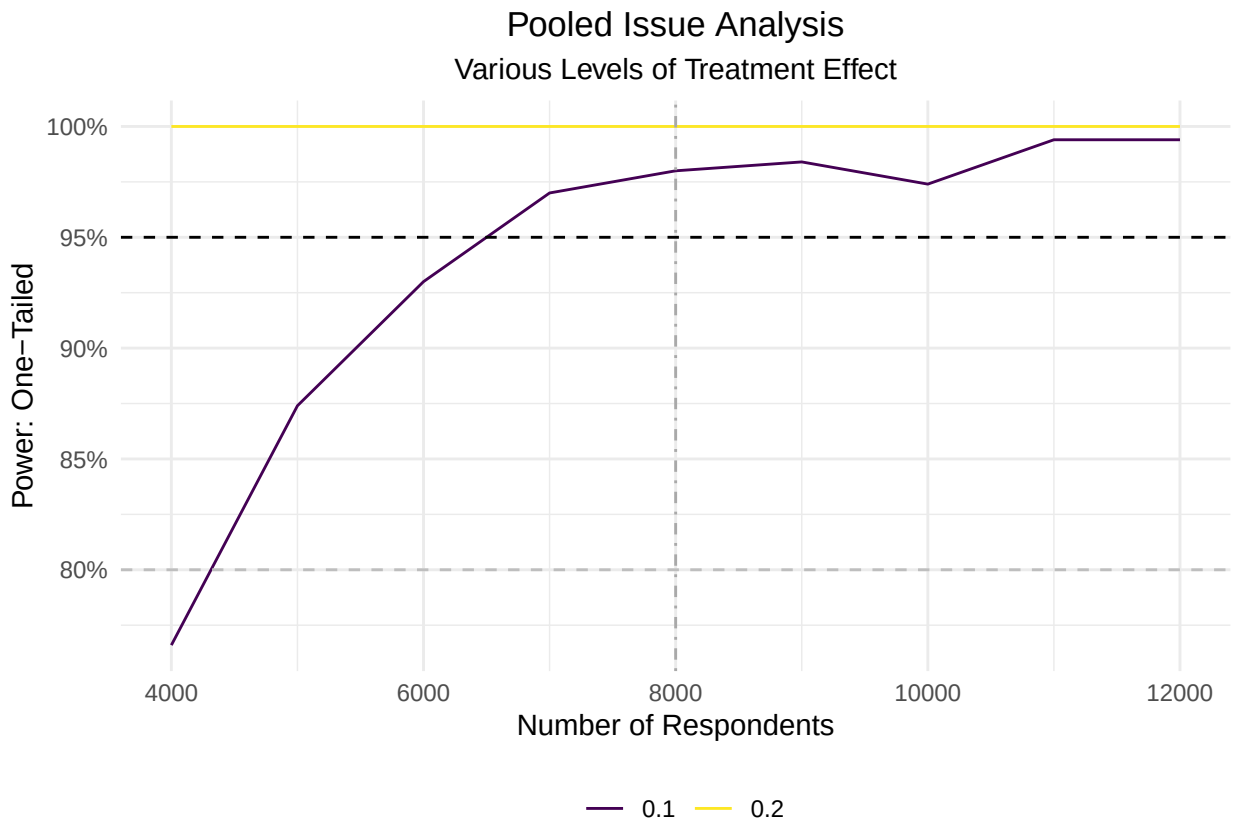
2.3 Power Analysis

As detailed in Section 4, we conduct an OLS regression for the four sentences. The manipulation, i.e. the masking condition vs. the political actor, as well as the specification of the sentence (i.e. *fully specified* or *underspecified*) as well as respondents ideological distance to the party and their levels of political knowledge are the main independent variables.

To calculate power for the hypotheses, the R package `DeclareDesign` is used (Blair et al. 2019). The effect sizes are between $b = 0.2$ and $b = 0.1$ – i.e. a small effect visualized by the purple and blue lines in Figure 1. The hypothesis are directional, Figure 1 therefore displays one-tailed tests with $\alpha = 0.05$. The power analysis shows that testing hypotheses 1 and 2 requires a sample size of $2,000$ participants, who all annotate the four

sentences, i.e. 8,000 participants in total, (x-axis) to reach very high levels of power, i.e. $> 95\%$, (indicated by the black dashed line in Figure 1).

Figure 1: Power Analysis



3 Measures

3.1 Dependent Variables

We rely on whether or not a party's (implied) stance is coded according to the party's position (H1 and H3) as well as whether or not the statement is coded as a stance at all (H3). For each issue, we ask the respondent `what is according to the sentence above the position of [ACTOR]?`, with the answer categories: `in favor`, `against`, `no stance`, `don't know`.

3.2 Treatment

Respondents are randomly assigned to either view a political party as an actor, or a masked condition, where they see `X` as an actor; simultaneously, respondents see either a fully specified sentence or a underspecified sentence, in which one needs additional information to interpret the positon on an actor. Table 1 gives an overview of the variations in treatment in the survey as well as their English translations.

3.3 Control Variables

As control variables, the following *demographics* are measured post-treatment: gender, age, education, geographical region, level of urbanness, employment, and income. For the analysis, only variables that are unbalanced over the experimental conditions will be included. Table 2 gives an overview of the questions asked in the survey as well as their English translations.

Table 1: Survey Questions - Experimental Conditions

Condition	Wording ENG	Wording NL
Specified	[PVV/X] says immigration should be made harder.	[PVV/X] zegt dat immigratie moeilijker gemaakt moet worden.
Specified	[GreenLeft/X] says nitrogen emissions need to be reduced.	[GroenLinks/X] zegt dat stikstofuitstoot meer tegengegaan moet worden.
Specified	[Labour Party/X] says tax rate should go up for highest earners.	[PvdA/X] zegt dat het belastingtarief voor de hoogste inkomens omhoog moet.
Specified	[Forum for Democracy/X] says that membership in the European Union has been especially bad for the Netherlands so far.	[Forum voor Democratie/X] zegt dat het lidmaatschap van de Europese Unie tot nu toe vooral slecht geweest voor Nederland is.
Underspecified	[PVV/X] says many immigrants are coming this way.	[PVV/X] zegt dat veel immigranten deze kant op komen.
Underspecified	[GreenLeft/X] says nitrogen policy must be different.	[GroenLinks/X] zegt dat het stikstofbeleid anders moet.
Underspecified	[Labour Party/X] says tax system must be changed.	[PvdA/X] zegt dat het belastingstelsel moet worden aangepast.
Underspecified	[Forum for Democracy/X] says the Netherlands should have a different role in the European Union.	[Forum voor Democratie/X] zegt dat Nederland een andere rol in de Europese Unie moet hebben.

- *Gender* is measured as `sex`. The answer categories are `Male` (value of `1`), `Female` (value of `0`), and `No answer` (value of `999`).
- *Age* is measured using 6 categories: `17 or younger`, `18--29`, `30--39`, `40--49`, `50--59`, `60--74`.
- *Education* is measured as the highest successfully completed level of education, recoded into four categories: `low`, `middle`, `high`, and `none`. I create dummy variables for each level of education with the lowest category as base category.
- *Employment* Respondents were asked which category of employment – `Full-time employed`, `Part-time employed`, `Entrepreneur`, `Unemployed and searching for a job`, `Unemployed and not searching for a job or incapacitated`, `Housewife/Househusband or else`, `Retired`, `Student or full-time education` – applied most to them.
- *Income* Respondents were questioned on their monthly income in bins of €500 – `€500 or less`, `€501-€1000`, `€1001-€1500`, `€1501-€2000`, `€2001-€2500`, `€2501-€3000`, `€3001-€3500`, `€3501-€4000`, `€4501-€7500`, `€7501 or more` – as well as giving them the options of `won't say` and `don't know`.
- *Geographical region* is measured using the *Nielsen districs*, dividing the Netherlands into 1) the 3 major cities plus suburbs, Amsterdam (plus Diemen, Ouder-Amstel, Landsmeer, Amstelveen), Rotterdam (plus Schiedam, Capelle aan den IJssel, Krimpen aan den IJssel, Nederlek, Ridderkerk, Barendrecht, Albrandswaard) and The Hague (plus Leidschendam, Voorburg, Rijswijk, Wassenaar, Wateringen); 2) West (Noord-Holland, Zuid-Holland and Utrecht (excluding the major cities and their suburbs; 3) North (Groningen, Friesland and Drenthe), 4) East (Overijssel, Gelderland and Flevoland); and South (Zeeland, Noord-Brabant and Limburg).

In addition, pre-treatment, respondents' ideological position, political knowledge, vote recall, and position on the issues migration, climate, tax, and EU are measured (see Table 3). Those variables will only be included in the analyses if balance checks indicate they are necessary. Moreover, the variables will be used to explore heterogeneous relationships.

- *Ideological position* is measured using an 11-point scale ranging from left (`0`) to right (`10`).
- *Vote Recall* Respondents were asked which party they voted for in the 2021 parliamentary elections.

Table 2: Survey Questions - Demographics

Variable	Wording ENG	Wording NL
gender (D1)	What is your gender?	Wat is uw geslacht?
age (D2)	What is your year of birth?	Wat is uw geboortejaar?
region (D3)	In which region do you live?	In welke regio woont u?
job (D5)	What applies most to you? Are you...	Wat is het meest op u van toepassing? Bent u...
income (D6)	Can you indicate which income groups your net monthly household income falls into?	Kunt u aangeven in welke inkomensgroepen uw netto maandelijks huishoudinkomen valt?
education (D7)	What is your highest level of education?	Wat is uw hoogst GENOTEN opleiding

The options were 1) all parties that were elected into parliament – `Bij1`, `BoerBurgerBeweging`, `CDA`, `ChristenUnie`, `D66`, `Denk`, `Forum for Democracy`, `JA21`, `GroenLinks`, `PvdA`, `Animal Rights Party`, `PVV`, `SGP`, `SP`, `VOLT`,`VVD`, `50Plus Party` – 2) another party; 3) blanco vote; and 4) a `Don't know` option.

- *Political knowledge* is measured with six items from the DPES.
- *Position on migration* is measured by asking people whether or not there are too many immigrants in the Netherlands using a 5-point Likert-scale (`fully disagree`, `disagree`, `neutral`, `agree`, `fully agree`).
- *Position on climate* is measured by asking people whether or not the climate crisis is exaggerated using a 5-point Likert-scale (`fully disagree`, `disagree`, `neutral`, `agree`, `fully agree`).
- *Position on tax* is measured by asking people whether or not the tax rate for the highest earners should go up using a 5-point Likert-scale (`fully disagree`, `disagree`, `neutral`, `agree`, `fully agree`).
- *Position on the EU* is measured by asking people whether or not membership in the European Union has been especially bad for the Netherlands so far using a 5-point Likert-scale (`fully disagree`, `disagree`, `neutral`, `agree`, `fully agree`).

3.4 Attention Check

We include one attention checks in the survey. This is asked just before respondents enter the round of the experimental treatments. The attention checks are taken from Berinsky, Margolis, and Sances (2014) and adapted to the German context by the authors. If a respondents fails the first attention check, a warning appears and the respondent can only continue with the survey once the respondent has correctly answered the question correctly.

Attention Check When a big news story breaks people often go online to get up-to-the-minute details on what is going on. We want to know which websites people trust to get this information. We also want to know if people are paying attention to the question. To show that you have read this much, please ignore the question and select Volkskrant and Metro as your two answers. When there is a big news story, which is the one news website you would visit first? (Please only choose one). Eight (Dutch) news outlets are provided to choose from. Respondents pass the attention check if they select *de Volkskrant* and *Metro*.

3.5 Exclusion Criteria

Participants are required to respond to each question.

4 Analysis

We test the hypotheses formulated in Section 1 by fitting linear multivariate regressions pooling the issues. In each model, we will estimate the coefficient for whether the actor in the sentence is masked or not, as well

Table 3: Survey Questions - PreTreatment Questions

Variable	Wording ENG	Wording NL
vote recall (PT1)	Which party did you vote for during the last parliamentary elections of March 2021?	Op welke politieke partij heeft u bij de vorige Tweede Kamerverkiezingen van maart 2021 gestemd?
introduction	To what extent do you agree with the following statements:	In hoeverre bent u het eens met onderstaande stellingen:
migration (PT2)	There are too many immigrants in the Netherlands.	Er zijn te veel immigranten in Nederland.
climate (PT3)	The climate crisis is exaggerated.	De klimaatcrisis wordt overdreven.
tax (PT4)	The tax rate for the highest earners should go up	Het belastingtarief voor de hoogste inkomens moet omhoog
EU (PT5)	The Netherlands should have a different role in the European Union.	Het lidmaatschap van de Europese Unie is tot nu toe vooral slecht geweest voor Nederland
RILE (PT6)	In politics there is often talk about left and right. Where would you put yourself on this scale?	In de politiek wordt vaak gesproken over links en rechts. Waar zou u uzelf op deze schaal plaatsen?
political-knowledge1 (PT7_1)	Which parties currently form the coalition government?	Welke partijen vormen op dit moment de coalitieregering?
political-knowledge2 (PT7_2)	Which of the following is the current Finance Minister in the Rutte cabinet?	Wie van de volgende personen is de huidige minister van Financiën in het kabinet Rutte?
political-knowledge3 (PT7_3)	Which political party was the second-largest party in the last Lower House elections on March 17?	Welke politieke partij was de op één na grootste partij tijdens de afgelopen Tweede Kamerverkiezingen van 17 maart?
political-knowledge4 (PT7_4)	By whom are members of the Senate elected?	Door wie worden de leden van de Eerste Kamer gekozen?
political-knowledge5 (PT7_5)	Who makes or is making new laws in the Netherlands?	Wie maakt of maken in Nederland nieuwe wetten?
political-knowledge5 (PT7_6)	If the government and the House of Representatives pass a law that is not in line with the Dutch Constitution, may the courts disapprove it?	Als de regering en de Tweede Kamer een wet aannemen die niet overeenkomt met de Nederlandse Grondwet, mag de rechter deze wet dan afkeuren?

as for whether the sentence is fully specified or underspecified. We will add ideological distance from the party mentioned as well as political knowledge as interactions with the experimental conditions. We will only add control variables in the analyses that are unbalanced, as explained in Section 4.1.

4.1 Balance Checks

We will conduct a balance test based on demographics (age, gender, education, geographical region, level of urbanness,employment, and income), vote choice in the 2021 parliamentary elections, ideological self-placement, political knowledge, and positions on the issues, using the `cobalt` R package (Greifer 2021). If the groups are unbalanced on one of these variables – i.e. standardized mean differences < 0.05 – we will add the covariates to the analyses. We will use the code below to conduct the balance tests (see here for the R script).

```
covs <- d %>%
  mutate(treatment = sentence) %>%
  select(sentence, D1:D7, PT1:PT&)

balanced <-bal.tab(treatment ~ factor(D1) + D2 +
                     factor(D3) + D4 + D5 +
                     factor(D6) + D7 + PT1 +
                     PT2 + PT3 + PT4 + PT5 +
                     PT6 + PT7,
                     data = covs,
          thresholds = c(m = 0.05))[[1]]
```

4.2 Hypothesis 1

To test the *Ideological Bias hypothesis*, we will conduct a multilevel model, with respondents clustered in issues, see code chunck below. Using the pooled data we will estimate a within groups fixed effects model. $\hat{YY}_{r,i,t}$ in Equation 1 denotes the evaluation of a stance by respondent r, during issue i and at experimental round t – ranging from round 1 to round 4. The standard errors are clustered at the individual level.

$$\hat{stance-correct}_{r,i,t} = \beta_0 + \beta_1 masked_{r,i,t} + \beta_2 specification_{r,i,t} + \beta_3 ideological-distance-to-party_{r,i,t} + \beta_4 political-knowled \quad (1)$$

We consider evidence for H1a when the effect of β_3 is postive and statistically significant. If the Average Marginal Effect of β_3 is positive and statistically siginificant when $\beta_2 = 1$ (i.e. condition is underspecidied), the second *Ideological Bias hypothesis* (H1b) will be confirmed.

```
h1a <- lmer(stance_correct ~ masked + specification +
              ideological_distance_to_party + political_knowledge +
              (1 | issue), data= d)

h1b <- lmer(stance_correct ~ masked +
              specification * ideological_distance_to_party +
              political_knowledge + (1 | issue), data= d)
```

4.3 Hypothesis 2

To test the *Overinterpretation Hypotheses*, we will conduct a multilevel model, with respondents clustered in issues, see code chunck below. Using the pooled data we will estimate a within groups fixed effects model. $\hat{YY}_{r,i,t}$ in Equation 2 denotes the evaluation of a stance by respondent r, during issue i and at experimental round t – ranging from round 1 to round 4. The standard errors are clustered at the individual level

$$\hat{stance-correct}_{r,i,t} = \beta_0 + \beta_1 masked_{r,i,t} + \beta_2 specification_{r,i,t} + \beta_3 heterogeneous_t reatmentr, i, t + \beta_4 specification * heterogen \quad (2)$$

When the Average Marginal Effect of β_3 ideological distance is positive and statistically siginificant when $\beta_2 = 1$ (i.e. condition is underspecidied), the *Ideological Overinterpretation hypothesis* (H2a) will be confirmed. When the Average Marginal Effect of β_3 political knowledge is positive and statistically siginificant when $\beta_2 = 1$ (i.e. condition is underspecidied), the *Political Knowledge Overinterpretation hypothesis* (H2b) will be confirmed.

```
h2a <- lmer(stance ~ masked + specification * ideological_distance_to_party +
              political_knowledge + (1 | issue), data= d)
```

```
h2b <- lmer(stance ~ masked +
              specification * ypolitical_knowledge +
              ideological_distance_to_party + (1 | issue), data= d)
```

4.4 Hypothesis 3

To test the *Masking Solution hypotheses*, we will conduct a multilevel model, with respondents clustered in issues, see code chunck below. Using the pooled data we will estimate a within groups fixed effects model. $\hat{YY}_{r,i,t}$ in Equation 3 denotes the evaluation of a stance by respondent r, during issue i and at experimental round t – ranging from round 1 to round 4. The standard errors are clustered at the individual level

$$\hat{stance} \; correct_{r,i,t} = \beta_0 + \beta_1 masked_{r,i,t} + \beta_2 specification_{r,i,t} + \beta_3 heterogeneous_t reatment r, i, t + \beta_4 masked * heterogeneous_t \quad (3)$$

When the Average Marginal Effect of β_3 ideological distance is negatiive and statistically siginificant when $\beta_2 = 1$ (i.e. condition is masked), the first *Masking Solution hypothesis* (H3a) will be confirmed. When the Average Marginal Effect of β_3 political knowledge is negative and statistically siginificant when $\beta_2 = 1$ (i.e. condition is masked), the second *Masking Solution hypothesis* (H3b) will be confirmed.

```
h3a <- lmer(stance ~ masked + specification * ideological_distance_to_party +
              political_knowledge + (1 | issue), data= d)

h3b <- lmer(stance ~ masked +
              specification * ypolitical_knowledge +
              ideological_distance_to_party + (1 | issue), data= d)
```

4.5 Statistical Significance

All the hypotheses are directional, and therefore all of the tests will be one-tailed. We will use an α-value of ‘0.05 as the value for statistical significance in all models above.

References

Berinsky, Adam J, Michele F Margolis, and Michael W Sances. 2014. "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys." *American Journal of Political Science* 58 (3): 739–53. https://doi.org/10.1111/ajps.12081.

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

Greifer, Noah. 2021. "Cobalt: Covariate Balance Tables and Plots. R Package Version 4.3.1." https://cran.r-project.org/web/packages/cobalt/index.html.