Good will hunting. Predicting response quality using motivation in longitudinal surveys
Valentin Brunel, Jean-Baptiste Portelli

To cite this version:
Valentin Brunel, Jean-Baptiste Portelli. Good will hunting. Predicting response quality using motivation in longitudinal surveys. European Survey Research Association, Jul 2019, Zagreb, Croatia. hal-02874071

HAL Id: hal-02874071
https://sciencespo.hal.science/hal-02874071
Submitted on 18 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ESRA 2019:
Good will hunting. Predicting response quality using motivation in longitudinal surveys

Valentin Brunel
Jean-Baptiste Portelli
Introduction

- Declining response rates and quality (Massey and Tourangeau 2013)
- What tools can we use to predict and target those challenges?
- Motivation: what is it?

---

Literature

- Most research on motivation is psychological
- Finding out what are the big drivers or sub-elements of motivation
- The objective in most cases is to maximize motivation
Literature

Most research on motivation is psychological

Finding out what are the big drivers or sub-elements of motivation

The objective in most cases is to maximize motivation

But this is not our case! We will simply define “motivation” as the reason why people say they participate in a survey (“functionalist approach”, Dillman 2000).

What we aim to do

- Find out to what extent motivation can explain data quality and attrition
- Use those results to maximize data quality and minimize attrition
- Plus another, more methodological issue: comparing open-ended questions and close-ended questions in producing an effective and useful classification of reasons to participate
Our data

ELIPSS Panel, two waves:

1. pilot study (2012, 1000 panelists)
2. refreshment (2016, 2500 panelists)

5 years of data: longitudinal vision, commentaries, paradata, full-text, etc.
Our variables

786 variables:

- Questions asked when joining the panel
- Paradata (response status, duration, number partial non-response) for all surveys
- Socio-demographics (each year)
- “Pratiques numériques” questions (each year)
- End-of-survey questions of these surveys: close-ended questions on motivation
- Other
Descriptive statistics

What is motivation and how correlated is it to other variables?
Descriptive statistics

Score constructed using:
- how much did you hesitate,
- did you subscribe yourself,
- did you check for additional information and where?
Other than the “quantitative” motivation score, we used a pair of questions asked to all panelists when entering: what is the reason why you chose to participate? Those questions had 6 possible answers:

- interest for research
- trust in institutions
- incentives (tablets or internet subscriptions)
- project’s originality
- other
Quali vs Quanti motivation

Another approach of qualitative motivation: textual analysis. Could you describe your first impressions when hearing about ELIPSS?

First results: very high similarity of vocabulary...

...which makes all analysis difficult.

<table>
<thead>
<tr>
<th>Word</th>
<th>Frequency</th>
<th>Part of Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>intéressant</td>
<td>990</td>
<td>adj</td>
</tr>
<tr>
<td>projet</td>
<td>621</td>
<td>nom</td>
</tr>
<tr>
<td>participer</td>
<td>398</td>
<td>ver</td>
</tr>
<tr>
<td>enquête</td>
<td>259</td>
<td>nom</td>
</tr>
<tr>
<td>trouver</td>
<td>248</td>
<td>ver</td>
</tr>
<tr>
<td>intéresser</td>
<td>243</td>
<td>ver</td>
</tr>
<tr>
<td>curiosité</td>
<td>239</td>
<td>nom</td>
</tr>
<tr>
<td>curieux</td>
<td>178</td>
<td>adj</td>
</tr>
<tr>
<td>tablette</td>
<td>161</td>
<td>nom</td>
</tr>
<tr>
<td>donner</td>
<td>149</td>
<td>ver</td>
</tr>
</tbody>
</table>
Quali vs Quanti motivation

However, we used Reinert (Alceste software) method, to identify classes of respondents, based on similarity of vocabulary.

Reinert method:

Descending hierarchical classification, which separates progressively texts depending on their similarity of vocabulary. Texts with the same words will surface in the same class.
Approaching motivation through all dimensions

Using a MCA

- close-ended answers on motivation
- textual analysis categories
- quantitative score
- having not answered surveys
- having started but not finished surveys
- difficulty of recruitment
- proportion of over-time survey answering
- proportion of item non-response in main surveys
Approaching motivation via different dimensions
Approaching motivation via different dimensions
Approaching motivation via different dimensions
Approaching motivation via different dimensions

The main result is the opposition between subjective motivation and objective measures of motivation.

The MCA seems to oppose 3 groups of people: clockwise, starting from top left:

- good-willing panelists
- hard-to-reach, not too interested from the start
- unconcerned, non respondents

The first axis is motivation enduring in time, the second motivation during recruitment.
Plan factoriel
Approaching motivation via different dimensions

One could here find other dimensions noted in the literature: extrinsic self (incentives), extrinsic other (recherche), intrinsic other (give opinion), intrinsic self (interest, curiosity).
Figure 1 Conceptual framework

- **Intrinsic motives**
  - Interest (Krosnick 1999; Sheehan & McMillan 1999; Van Kenhove et al. 2002)
  - Enjoyment (Porst & Von Briel 1995)
  - Curiosity (Porst & Von Briel 1995)

- **Extrinsic motives**
  - Incentives (Gortz 2004; Groves et al. 2000)

- **Focus**
  - Helping (Dillman 2000; Evangelista et al. 1999; Groves et al. 1992; Porst & Von Briel 1995)
  - Give opinion (Dillman 2000; Evangelista et al. 1999; Groves et al. 1992)

- **Self**

- **Other**

- **Need for recognition** (Dillman 2000; Toepfer 1930; Warwick & Lininger 1975)
A statistical model to predict non response

- Model to predict attrition under-performed (16.5% bar of well-classed panelists)

- Thus we turned towards predicting non-response to at least one survey as our variable of interest.
Our model

Variables selected:
- durationsum2_cl: response time to an inquiry
- ea16_A2A_REC_cl): age divided in 5 slices
- housing16: occupancy status of the dwelling (owner or tenant)
- qmotiv2: qualitative initial motivation variable in 4 groups (Research, Originality of the project, Tablet and subscription, other)
- score_cl: initial motivation score in slices
- difficulty_recruitment: difficult to find during recruitment, measured by survey institute
Coefficients:  

| Term                          | Estimate | Std. Error | z value | Pr(>|z|) |
|-------------------------------|----------|------------|---------|----------|
| (Intercept)                   | 18.1974  | 3956.1805  | 0.005   | 0.99633  |
| score_cl[1,2]                 | -0.38742 | 0.44544    | -0.870  | 0.384440 |
| score_cl[2,6]                 | -0.44104 | 0.33451    | -1.318  | 0.187341 |
| score_cl[6,8]                 | -0.37790 | 0.33641    | -1.123  | 0.261291 |
| score_cl[8,11]                | -0.55896 | 0.33864    | -1.651  | 0.098823 |
| ea16_A2A_REC_cl[7,9]          | -0.28524 | 0.14118    | -2.020  | 0.043346 |
| ea16_A2A_REC_cl[9,11]         | -0.51901 | 0.15435    | -3.363  | 0.000772 *** |
| ea16_A2A_REC_cl[11,99]        | -0.78053 | 0.23635    | -3.302  | 0.000959 *** |
| durationsum2_cl[0.25,0.35]    | -0.18228 | 0.18209    | -1.001  | 0.316795 |
| durationsum2_cl[0.35,0.4]     | -1.17636 | 0.18044    | -6.519  | 7.06e-11 *** |
| durationsum2_cl[0.4,0.45]     | -2.16779 | 0.17701    | -12.247 | < 2e-16 *** |
| durationsum2_cl[0.45,1]       | -3.76696 | 0.24150    | -15.598 | < 2e-16 *** |
| qmotiv2Originalité du projet  | -0.68006 | 0.25698    | -2.646  | 0.008137 ** |
| qmotiv2Recherche              | -0.68388 | 0.23604    | -2.897  | 0.003763 ** |
| qmotiv2Tablette et ab         | -0.59665 | 0.24347    | -2.451  | 0.014262 * |
| difficult_recrutement         | -0.70899 | 0.15675    | -4.523  | 6.09e-06 *** |
| logement16Accédant à la propriété | 0.09051 | 0.13705 | 0.660  | 0.508974 |
| logement16Locataire           | 0.35847  | 0.13614    | 2.633   | 0.008462 ** |
| logement16Gratuit             | 0.28292  | 0.25257    | 1.120   | 0.262640 |
Our model
Rate of well ranked = 67%

Rate of well ranked = 63%
Our model

Although the model described above is not perfect, we can present the typical profile of a panelist's behavior with respect to his or her likelihood of not responding to an investigation can be identified. These criteria are:

- a low initial motivation score
- An initial motivation to acquire tablets and subscription
- Difficulties in contact during recruitment
- a young age
- tenant status
Conclusion

- Motivation, even initial, can be useful in predicting and preventing non-response. Surprising result: data collected sometimes seven years earlier.
- Paradata as a great indicator of motivation and a very useful tool: looking for other variables...
- Usg those elements could prove interesting in organising panel calls and stimuli. Focus on panelists with more risk of leaving/not answering.

- Textual analysis of motivation, even with a very homogeneous corpus, looks promising.