Reliability testing of coded election manifesto data in the DIPAC-project

Karl Loxbo, Sofie Blombäck, Jonas Hinnfors, Marie Demker and Magnus Hagevi
Contents

Contents ........................................................................................................................................ 2

Introduction .................................................................................................................................. 3

Table 1. Dimensions and sub-dimensions in election manifestos .................................................. 4

The coding process .......................................................................................................................... 4

Reliability analysis .......................................................................................................................... 4

Results of the reliability analysis ..................................................................................................... 5

Figure 1. Reliability test for parties’ saliency and position on the Left-Right dimension and subdimensions .................................................................................................................................. 6

Figure 2. Reliability test for parties’ saliency and position on the GAL-TAN dimension and subdimensions .................................................................................................................................. 6

Figure 3. Reliability test for parties’ saliency and position on remaining dimensions. .................................................................................................................................. 7

Table 2. Dimensions and sub-dimensions in the DIPAC and the CMP/MRG .................................................................................................................................. 8

Figure 4. Comparison of reliability scores in the DIPAC and the CMP/MRG on the Left-Right dimension and subdimensions .................................................................................................................................. 10

Figure 5. Comparison of reliability scores in the DIPAC and the CMP/MRG on the GAL-TAN dimension and subdimensions .................................................................................................................................. 10

Figure 6. Comparison of reliability scores in the DIPAC and the CMP/MRG on remaining dimension. .................................................................................................................................. 11

References ......................................................................................................................................... 11
Introduction
The data described in this report were collected for the DIPAC-project. The project’s aim was to closer investigate the proposed crisis of parties and party systems in Western Europe during the last decades. To fulfil this aim, one of our main goals was to collect electoral manifests from all parties in the respective party system for all selected years, and code their ideological contents. The main reason for our choice to code manifests was to provide a more detailed and nuanced foundation for analyzing ideological polarization and ideological change in Western European party system than the saliency-oriented perspective that has hitherto dominated manifesto-research.

The Comparative Manifesto Project (CMP), nowadays referred to as the Manifesto Research Group (MRG), provide comprehensive data on parties’ positioning in manifests that is comparable across time and space. Through the coding of quasi-sentences – that is, sections in the text that contain ‘the verbal expression of one political idea or issue’ (Budge et al. 2001: 217) these data provide information on the proportion of parties’ manifests (in relation to the whole manifesto) that is devoted to a certain issue. If a manifesto makes frequent references to a certain policy issue, the party behind it is believed to view it as a salient feature in its electoral strategy. In contrast to this approach, the coders in the DIPAC-project have focused on coding ideological positions as such – that is, ideological visions, frames or worldviews that give voters an idea of how the party would like to see society develop and why. That means that DIPAC-project deviates from the saliency-oriented approach of the CMP/MRG and instead focuses on the manifesto as a whole. The coding is thus interpretative, with the coder making a judgment for each ideological sub-dimension. This also means that not every sentence is coded. Instead, the coder has identified typical quotes illustrating the party’s reasoning on each dimension – along with any ‘problematic’ statements that might make the party’s position unclear.

The DIPAC-project codes parties’ ideologies into two dimensions, each with several sub-dimensions. These are listed in Table 1.

---

1 Decreased Ideological Polarisation and Conflict in Western Europe. The project was funded by a grant from the Swedish Research Council.
Table 1. Dimensions and sub-dimensions in election manifestos

<table>
<thead>
<tr>
<th>Left-Right</th>
<th>GAL-TAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>state-market</td>
<td>environment – economic growth</td>
</tr>
<tr>
<td>public-private welfare provision</td>
<td>civil liberties – law &amp; order</td>
</tr>
<tr>
<td>high – low redistribution</td>
<td>Individual liberty – social authoritarianism</td>
</tr>
<tr>
<td></td>
<td>Multiculturalism – national unity</td>
</tr>
<tr>
<td></td>
<td>Cosmopolitanism - Nationalism</td>
</tr>
<tr>
<td></td>
<td>EU positive – EU negative</td>
</tr>
</tbody>
</table>

The focus of the coding is always on the ideological reasoning of the parties, in that there are no policy proposals that automatically place the party in a particular category on a particular dimension.

The coding process

The coding was carried out by a small team of coders recruited among MA students and recent MA graduates at the department of political science at the University of Gothenburg, led by one of the project researchers. All manifestos from the same party system were coded by the same coder, to ensure as much within country consistency as possible. At the same time, the coders worked in close proximity to each other, and were encouraged to discuss their coding to enable them to keep the whole project team ‘on the same page’. Each coder was given instructions in the coding manual and the NVivo software used, and briefed on the aim of the coding – that is to achieve a qualitative overall assessment of the party’s stance on each ideological sub-dimension. The coders read through each manifesto, flagging sections of text that were relevant for identifying an ideological position on one or more sub-dimensions. The coder then summarized each manifesto into a coding sheet, indicating saliency, position and potential blurriness for each dimension, as well as the main enemy, main problem and focus on ideology or competence.

Reliability analysis

To ensure that our data are stable and reliable, a new coder coded approximately 19 percent of the material. Krippendorf’s Alpha (Krippendorf 2004), was calculated for the saliency and position for each of the sub-dimensions. Data for Germany and the United Kingdom, which constitutes 19 per cent of the current database (without the Netherlands and all years for Belgium), was coded a second

---

2 The four coders were Sofie Blombäck, Lena Caspers, Laura Lungu and Ruben Dielman. At the initial stages of the project Sara van der Meiden coded some manifestos while contributing to the development of the coding scheme. All these manifestos were recoded by a different coder one the coding scheme was finalized.
time by a new coder. The recoded data was subsequently compared with the
original data, column by column, in a new database.

Krippendorf’s Alpha not only considers the observed, but also the expected,
disagreement between two coders. The observed disagreement indicates the
percentage of mismatches between coders in values attributed to the same units,
whereas the expected disagreement is a probability measure that coders code a
unit correctly just by chance. If there are more different values, the chance to
code rightly by random picking is smaller. That is to say, that the probability for
picking the correct alternative randomly is significantly lower if scales are ordinal
and not nominal. Krippendorf’s Alpha is calculated in the following way:

$$\text{Krippendorf’s Alpha} = 1 - \frac{\text{Observed disagreement}}{\text{Expected disagreement}}$$

Krippendorf’s Alpha is basically a measure of correlation, ranging from 0
to 1, between two variables (the same variable coded twice) weighed by
the probability that correct coding is random. Kalpha=.80 is often brought
forward as the norm for a good reliability test, with a minimum of .67 or
even .60 (when it is that low, you might give some specific information
why this is low and why you still choose to accept this variable in your
analysis). However, a .80 or higher Kalpha should not always be
considered satisfactory. When the variable (in the eyes of the researcher) is
extremely easy to code one might raise the standards a little bit.

To implement the reliability test, the code for KAPHLP was first written
into SPSS (syntax) (http://www.afhayes.com/spss-sas-and-mplus-macros-
and-code.html). However, since STATA allows for the computation of
ordinal scales, we chose to implement our test using this program.

**Results of the reliability analysis**

The measures of Krippendorf’s Alpha, for each ideological sub-dimension
as well as for the Left-Right and GAL-TAN scales in total, are presented
in Figures 1 and 2. The bars in the figures display Krippendorf’s Alpha if
the measures are assumed to be either on an ordinal or on a nominal scale.
Given that our aim is to measure ideological distance on equivalent scales,
the most relevant measure of reliability is here the bars which display
Krippendorf’s Alpha if scales are assumed to be ordinal.
Figure 1. Reliability test for parties’ saliency and position on the Left-Right dimension and subdimensions.

Figure 2. Reliability test for parties’ saliency and position on the GAL-TAN dimension and subdimensions.
Figure 3. Reliability test for parties’ saliency and position on remaining dimensions.

When reviewing the graphs, we see that both ideological dimensions and almost all sub-dimensions receive Krippendorf’s Alpha scores that are close to, and mostly well above, the 0.8 standard. Yet, some of the saliency measures for the GAL-TAN sub-dimensions – especially Multiculturalism-Nationalism (0.51-0.64) – have lower scores and are consequently less reliable. Thus, the coding of saliency for some of the GAL-TAN sub-dimensions should be interpreted with caution. However, it is important to point out that all of our measures of parties’ ideological position receive scores around 0.8 and mostly above. These measures can, therefore, be considered highly reliable.

Moreover, we compare our Krippendorf’s Alpha scores with the same measures in the CMP: Building on existing codes for the same parties, countries, and years we have constructed the same ideological subdimensions in the CMP that are displayed in Table 2.
<table>
<thead>
<tr>
<th>Table 2. Dimensions and sub-dimensions in the DIPAC and the CMP/MRG</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DIPAC operationalisations</strong></td>
</tr>
<tr>
<td><strong>Left-right scale</strong></td>
</tr>
<tr>
<td>State-market</td>
</tr>
<tr>
<td>regulating free market, state’s role in economy, size of state intervention</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Public-Private welfare provision</strong></td>
</tr>
<tr>
<td>Who provides public goods, private or public solutions to eg medical insurance etc</td>
</tr>
<tr>
<td><strong>High – low redistribution</strong></td>
</tr>
<tr>
<td>Redistribution: between rich and poor, taxation etc</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Civil liberties – law and order</strong></td>
</tr>
<tr>
<td>Civil rights -Law and order: GAL promotes high protection of civil rights, rule of law, freedom from government surveillance, no mandatory voting etc TAN s focus on security, tough on crime, more resources for law enforcement, national security</td>
</tr>
<tr>
<td>Individual liberty – social authoritarianism</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Multiculture – National unity</td>
</tr>
<tr>
<td>GAL- protection of minorities, open migration policies, belief that several cultures can coexist, TAN – assimilation, restrictive migration policies, protection of traditional way of life</td>
</tr>
<tr>
<td>Cosmopolitanism – Nationalism</td>
</tr>
<tr>
<td>GAL- cosmopolitan, anti-military, EU-positive TAN: more resources for military, promoting national sovereignty</td>
</tr>
<tr>
<td>Pro EU – Anti EU</td>
</tr>
<tr>
<td>Ideology-competence</td>
</tr>
</tbody>
</table>

Note: per… refer to the name of the coding in the CMP/MRG and p. refers to the page number in the CMP/MRG codebook.

When reviewing Table 2, we see that we have measures on our two ideological dimensions and all of our sub-dimensions. Although the coding in the CMP/MRG concern h quasi-sentences in order to assess the saliency of issues, while we code saliency and position on ideological dimensions and sub-dimensions as a whole, the Krippendorf’s Alpha measures of reliability in the CMP/MRG are compared to the ones we report. Comparisons between scores from the DIPAC reliability test and the CMP/MRG reliability test are displayed in Figures 4-6.
Figure 4. Comparison of reliability scores in the DIPAC and the CMP/MRG on the Left-Right dimension and subdimensions.

Figure 5. Comparison of reliability scores in the DIPAC and the CMP/MRG on the GAL-TAN dimension and subdimensions.
The comparison between the DIPAC reliability test the CMP/MRG test leads to the following two conclusions about the reliability of our data. First, while the CMP/MRG shows consistent scores across all sub-dimensions (around 0.78) the reliability of our data is in fact far better than that of the CMP/MRG with regard to Left-right dimension and sub-dimensions. Second, the DIPAC measures of saliency is clearly less reliable than the CMP/MRG with regard to all GAL-TAN sub-dimensions. However, also with regard to these sub-dimensions the reliability of our measures of ideological position are clearly better than in the CMP.

In sum, we conclude that the DIPAC coding of ideological position has yielded highly reliable measures, whereas the coding of saliency should be interpreted with some caution with regard to the GAL-TAN dimension and sub-dimensions.

References
