Spatial models in voting advice applications

Otjes, Simon; Louwerse, Tom

Published in:
Electoral Studies

DOI:
10.1016/j.electstud.2014.04.004

Publication date:
2014

Electoral Studies 36 (2014) 263–271

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Spatial models in voting advice applications

Simon Otjes [a,*], Tom Louwerse [b]

[a] Documentation Centre Dutch Political Parties, Groningen University, Broerstraat 4, Groningen, The Netherlands
[b] Department of Political Science, Trinity College Dublin, 3 College Green, Dublin, Ireland

ARTICLE INFO

Article history:
Received 18 December 2012
Received in revised form 19 December 2013
Accepted 1 April 2014
Available online 16 April 2014

Keywords:
Voting advice application
Spatial model
Political party
E-democracy
Electoral behaviour

ABSTRACT

We examine the use of spatial models in Voting Advice Applications. A successful branch of VAAs uses multidimensional models of the political space to inform users of their policy match with political parties. Creators of these VAAs offer only a very generalized justification for the choice of their underlying spatial model. We examine whether these spatial representations offer a valid depiction of policy differences between parties. We compare the spatial models from the available national 'electoral compasses', which include established democracies as well as semi-authoritarian and transitional systems. We find that the two-dimensional 'one-size fits all' model that is used in all of these countries fails to accommodate significant variation between party systems.

1. Introduction

Studies have shown that Voting Advice Applications (VAAs) play a substantial role in shaping voters' perceptions of the positions of political parties (see other articles in this symposium). Voters and politicians take the information these tools provide seriously. This places a strong obligation of VAA creators to carefully examine their design choices. Research has shown that differences in statement selection and the method to calculate results have a large impact on the advice users receive (Louwerse and Rosema, 2013; Lefevere and Walgrave, 2014). This makes it pivotal to study the reliability of VAA design. This article examines the spatial models underlying VAAs.

Some VAAs provide advice by asking respondents' opinions on a number of statements and calculating the number of matches between a user and a number of parties. More recently, VAAs have started to use spatial modelling techniques derived from economic and political science in an attempt to use "scientifically approved methods" to place political parties (cited in Fossen and Anderson, 2014). Examples of these VAAs include the series of 'electoral compasses' (such as *Kieskompas*, *Bússola Eleitoral* and *Valkompass*), the *EU Profiler* and *Smartvote*. These VAAs do not just inform citizens about their policy match with each party, but also provide insight into the positions of political parties in (a representation of) the political space. These tools provide citizens the option of exploring their agreement with parties more precisely but primarily presents the 'match' as closest party in the spatial models.[1] The creators of these VAAs usually provide only a very generalized justification for the choice of their underlying spatial model. In most cases, designers do not base their selection of which dimensions to include in their spatial model (such as left–right and progressive–conservative) on empirical patterns in the data (Costa Lobo et al., 2010).

This article aims to test whether the spatial models used in VAAs meet basic scientific standards. We specifically examine *whether the spatial representations offer a valid depiction of policy differences between parties*. We examine

* Corresponding author. Tel.: +31 (0) 50 3636830.
E-mail addresses: s.p.otjes@rug.nl, simon@simonotjes.nl (S. Otjes), tom.louwerse@tcd.ie (T. Louwerse).

[1] In all of the VAAs we study the 'compass' is the only or the primary way in which the voting advice is presented.

http://dx.doi.org/10.1016/j.electstud.2014.04.004
0261-3794/

the spatial models of ten national VAAs, both in developed democracies such as Canada, as well as transitional democracies such as Tunisia.

The article is organized as follows. First, we discuss the use of spatial models in VAAs and the theoretical grounds for their usage. Second, we introduce our case selection of ten VAAs from the 'electoral compass' stable and outline their design. We examine the scalability of the various electoral compass dimensions and explore the inductive creation of models that fit parties' positions better. Results show that the scalability of the left–right scale is moderate to strong in most countries, but the scalability of the progressive–conservative scale is poor. Moreover, there is only a moderate fit between the model used in the electoral compasses and inductive models derived from parties' actual answers to the statements. This has significant implications for the status of the advice these applications provide, as the models in which voters and parties are placed are poorly specified and therefore it is unclear whether they are substantively meaningful.

2. The use of spatial models in VAAs

VAAs collect and represent party positions in different ways. One group of VAAs, including the *Walh-O-Mat* asks parties to indicate their position on about 30 statements. The VAA compares the answers of the users to the parties' responses and counts the number of 'matches' between the answers of the user and the parties. These VAAs do not explicitly use a spatial model, although one could argue that they implicitly use a high-dimensional conception of the political space (Louwerse and Rosema, 2013).

The use of spatial modelling techniques in political science has inspired the development of a different type of VAA, which explicitly uses a spatial model to position parties and voters. The Dutch *Kieskompas* was the first to use such a model, but the method has been copied in many other countries as well as the EU Profiler, a pan-European VAA for the European Parliament elections of 2009.[2] The underlying idea of these VAAs is that the political competition is structured by a limited set of dimensions (Krouwel and Fiers, 2008). Parties are positioned on thirty issue statements, using a five point scale ranging from fully disagree to fully agree. Each statement is assigned to one of the dimensions. A party position on a dimension is the average of their answers to the statements relating to that dimension.[3] This results in a two-dimensional (sometimes three-dimensional) representation of that is claimed to represent 'the political space'. The result screen of these VAAs displays this spatial model, indicating the estimated position of the user. The vote advice is implicit – the message to the user is: you have more in common with parties that are closer to you.

While two-dimensional models have been used most often in VAAs, some VAAs have used higher-dimensional representations. One example is the Swiss *Smartvote*, which displays the match between a user and a candidate for parliament in a so-called 'spider diagram'. This diagram shows the overlap between the user and the politician on eight dimensions. Here, the spatial model is used as visual tool rather than a way of calculating the 'match' between users and parties – *Smartvote* uses the Euclidean distance measure for the latter.

3. Spatial modelling

The spatial representations used in VAAs originate from a long-standing tradition of spatial modelling in the political science literature (Downs, 1957). Sometimes described as the "workhorse theory" of the study of politics (Cox, 2001: 189), spatial models have been used in electoral research, legislative studies, political communication and many other fields. The underlying assumption of spatial modelling is that parties do not pick their issue positions at random, but that there is an underlying, low-dimensional space which structures party positions. The dimensions of the spaces of party competition and the party positions cannot be directly observed. The goal of spatial modelling and scaling analysis is to move from observed manifest items to unobservable latent dimensions. Spatial modelling is a method of data reduction: one moves from an overwhelming number of observations to a limited number dimensions. This makes it easier to understand the differences between parties. Moving from observations to models is a process of creation: "[t]here is no true but unobserved "dimensionality" of any policy space waiting to be revealed by careful analysis (Benoit and Laver, 2006: 110). Even the godfather of dimensional scaling methods, Coombs (1964) emphasised that a scientist *creates* a model by selecting which observations he will use, how they will be classified and how they will be aggregated.

Creating a model for use in voting advice applications is relatively difficult, because one starts with a large number of statements rather than a limited set of policy dimensions. The question how many dimensions are necessary to capture the essential differences between political actors is central to most scaling exercises (Benoit and Laver, 2012). Unless party positions on specific issues correlate perfectly, one will always lose some information when moving from a high-dimensional to a low-dimensional model. What degree of data reduction is acceptable? Benoit and Laver (2012) argue that this choice should be based on both prior knowledge of the political system as well as patterns that exist in the data. Conventional assumptions about which dimensions structure the political space should be tested to see whether the model actually fits the data.

Most scaling methods are stochastic in nature: moving along the dimension means that the *chance* of a positive response from the subject increases (Niemöller and Van Schuur, 1983). It is crucially important to test whether this is actually the case: whether the items one believes to form a consistent scale actually do so. If this does not hold the scale misrepresents the differences between political actors. We take the perspective of consistency: to what extent do the different items measure the same underlying

[2] The EU Profiler uses a two-dimensional model alongside the other a matchlist and spider diagram.

[3] Note that the direction of some items is reversed, e.g. disagreement with a left-wing proposal implies a right-wing position and vice versa.

concept? This question can be approached from two sides: from the voter perspective or 'demand side' or from the party perspective or 'supply side'.

A model is internally valid from the 'demand side' perspective if it conforms to the underlying patterns of voters' preferences. We know only of a limited number of studies investigating this. One example concerns the Dutch VAA *Kieskompas*; whether the spatial model in this VAA is indeed an accurate reflection of how voters perceive the political competition has been tested by Kleinnijenhuis and Krouwel (2008); discussed in Van Kersbergen and Krouwel (2008). They constructed a model of the voter space based on the answers of a sample of the 1.7 million Kieskompas users. Using factor analysis, they uncovered the latent factors that structure the answers of the respondents.[4] They expected to find four orthogonal dimensions: two relating to economic left–right distinction and two relating to GAL–TAN distinction, but instead found six (Kleinnijenhuis and Krouwel, 2008). Only one of the uncovered dimensions exclusively concerns social-economic issues. Two dimensions mix issues from the social-economic left–right and GAL/TAN dimension. Three dimensions relate to different aspects of the GAL–TAN dimension. Van Kersbergen and Krouwel (2008) conclude: "the extracted factors do not entirely confirm the existence of a bipolar left–right dimension and a bipolar GAL–TAN dimension, but they do indicate a pattern of factors largely coherent with the assumed main dimensions of political competition in the Netherlands". To us, this seems an optimistic reading of their results: Kieskompas' two-dimensional representation of the 2006 Dutch political space does not relate strongly to an inductive model of the political space based on voters' positions. Although this analysis concerns only a single VAA, difficulties in modelling voters' answers to VAA statements have also been noted by Louwerse and Rosema (2013) and Wheatley (2012). This indeed confirms Converse's (1964) observation that the ideological constraint among the mass public is limited. It is thus very difficult to construct a model that fits very well with the 'demand side' of politics.

An alternative way of looking at the spatial models used in VAAs is that they offer an accurate representation of the 'supply side', that is the structure of party competition. While it is difficult to summarize policy views of a large number of citizens in a low dimensional space, it seems more feasible to construct a model that captures the policy differences between political parties. We evaluate this question by looking at three aspects: the extent to which the deductive scales from the different electoral compasses are consistent, the dimensionality of the political space based on an inductive analysis of party positions and the similarity of the deductive and inductive spaces.

[4] One may question whether factor analysis is appropriate for this enterprise. Kleinnijenhuis and Krouwel (2008) themselves cite a paper titled "Why factor analysis often is the incorrect model for analysing bipolar concepts, and what model to use instead" (Van Schuur and Kiers, 1994).

4. Research design

One prominent example of a VAA that uses spatial models is the *Kieskompas*. Its designer has, in cooperation with country specialists, developed 'electoral compasses' for more than ten countries.[5] These VAAs present an excellent set of cases for an analysis of spatial modelling in VAAs, because they use similar methods and a similar spatial model across countries. This allows us to examine whether the spatial model used in these VAAs is a valid representation of the party political competition in each country. We examine the spatial models in all national electoral compasses for which data was available: Belgium, Portugal, the Netherlands, Sweden, Canada, Turkey, Tunisia, Morocco, Egypt and France[6] (see Table 1). The selection includes transitional democracies (Tunisia and Egypt), a semi-authoritarian system (Morocco), as well as seven established democracies. Note that, despite the potentially large differences between democratic, transitional and semi-authoritarian systems, the VAA designers use the same two-dimensional model (left–right and progressive–conservative) in each of these cases. This offers an excellent test for the question whether the two-dimensional electoral compass model holds across borders.

The political spaces used in the electoral compasses were designed in the following way. In most cases the creators used a two-dimensional model with a social economic left–right dimension and a cultural dimension concerning religious traditionalism, support for the European Union, environmentalism and migration (Costa Lobo et al., 2010; Kleinnijenhuis et al., 2007; Van Kersbergen and Krouwel, 2008). The latter dimension is usually labelled 'progressive versus conservative' or 'Green Alternative Libertarian versus Traditional Authoritarian Nationalist' (cf. Hooghe et al., 2002). The selection of statements in the electoral compasses was based on the relevance of political issues, such as migration, the environment or taxes (Costa Lobo et al., 2010). These issues were selected on basis of expert judgement, opinion polls and manifesto coding (Kleinnijenhuis et al., 2007). For each issue, three to five statements were selected. Parties' positions on the statements are determined on the basis of expert judgements, text analysis and party self-placement (Kleinnijenhuis et al., 2007). The last step in the design of the electoral compasses is to assign statements to either of the two dimensions "based on an assessment as to which dimension an issue belongs as well as with respect to the question as to which side of a dimension a specific issue belongs." (Van Kersbergen and Krouwel, 2008). This assignment of

[5] We use the term 'electoral compass' as a generic term for the VAAs that have been developed by Kieskompas in corroboration with national experts and organizations. We limit our analysis to these VAAs, because it allows us to examine the extent to which their two-dimensional spatial model travels between countries.

[6] We limited the selection to one VAA per country. Unfortunately, the data for the Israeli electoral compass was no longer publicly available. The EU Profiler is discussed extensively by Louwerse and Otjes (2012). The USA electoral compass was not developed for the American public, but rather for the Dutch public; moreover, only the data for the two main candidates was still available, which makes it problematic to evaluate scalability.

Table 1
Cases included in the analysis.

VAA name	Country	Year	Kind of election	Type of democracy
Kieskompas	Belgium	2006	Parliamentary	Parliamentary
Bússola Eleitoral	Portugal	2009	Parliamentary	Parliamentary
Kieskompas	The Netherlands	2010	Parliamentary	Parliamentary
Valkompass	Sweden	2010	Parliamentary	Parliamentary
Vote Compass	Canada	2011	Parliamentary	Parliamentary
Genel Seçimleri	Turkey	2011	Parliamentary	Parliamentary
Boussole Electorale	Tunisia	2011	Parliamentary	Transition
Boussole Electorale	Morocco	2011	Parliamentary	Semi-authoritarian
Vote Compass	Egypt	2011	Parliamentary	Transition
Boussole Présidentielle	France	2012	Presidential	Semi-Presidential

statements to the dimensions is based on substantive considerations only, not on structures underlying parties' or voters' positions on the statements (Kleinnijenhuis and Krouwel, 2008).

In analysing the scalability of the electoral compass models, we use Mokken scaling analysis (Mokken and Lewis, 1982). Mokken scaling is a probabilistic approach to scaling, which assumes that the probability of a positive response to an item increases with increasing subject value (Van Schuur, 2003). It was developed for exam questions: The probability that a student gets each of the exam questions right increases with his exam grade. A student who gets the difficult questions right, should also have a high probability to answer the easy questions correctly. For party positioning, this means that the probability that a party agrees with a right-wing (left-wing) proposal increases if a party is more right-wing (left-wing). Similarly, if a party agrees with a strongly right-wing proposal, it should tend to support a moderately right-wing proposal as well.

Mokken scaling can be used in a confirmatory and an exploratory way. First, we aim to confirm the strength of the scales used in the various electoral compasses.[7] Loevinger's coefficient of homogeneity (H) is calculated to this end. This coefficient equals 1 when a set of items form a perfect Mokken scale and 0 when there is no statistical association between users' answers on the items. Mokken (1971) suggests the following benchmarks for Loevinger's H: below 0.3 (very poor), below 0.4 (poor), below 0.5 (medium), below 0.6 (strong), beyond 0.6 (very strong). One can calculate the homogeneity coefficient for a scale as a whole, but also for individual items or a specific pair of items.

Next, we use Mokken Scaling in an exploratory way: we try to find strong scales based on parties' answers on the statements. This 'automated item selection procedure' (AISP) works as follows (Mokken, 1971). First, one selects the two items with the highest H-coefficient; this is the starting pair for the first dimension. Then, one adds the item, which results in the highest H-coefficient when added to the starting pair. This step is repeated until no items remain or the H-coefficient would drop below a certain minimum (in our case 0.3). If more than one item remains, the AISP will try to create a second dimension using the same procedure.

One complication in our case is that the AISP assumes that all items are coded in the same direction (Van der Ark, 2007). This would be the case when one would, for example, construct a dimension of prejudice and the items measure the extent to which one dislikes certain minorities. A higher score on each of these items implies a higher level of prejudice. The statements we use are different in the sense that we do want to assume before our analysis whether agreeing with a certain statement is left-wing or right-wing (or progressive or conservative). Therefore we allow the 'direction' of a statement to change. For example, we do not assume that support for a directly elected mayor is a progressive or left-wing stance. In fact, in the Dutch case we find that right-wing parties are more likely to support direct election of mayors. Thus, the exploratory procedure we follow treats it as an empirical question what the appropriate number of dimensions is and how the statements relate to these dimensions.

Like all data-reduction methods the AISP-procedure seeks to make a model with a minimum number of dimensions, while still correctly representing the underlying party positions. We do not claim that the models we present prove that the 'true' political space has one, two or three dimensions. Our goal with the AISP is threefold: first, to assess whether we can inductively discover whether or not there are models that actually fit the party positions on many items. Second, to show that some models improve in scalability if items are 'flipped' in direction from the original models. Third, to demonstrate that specific items

Table 2
Scalability of the electoral compasses scales.

Country	Left–right	Progressive–conservative
Belgium	0.515	0.259
Canada	0.665	0.354
Egypt	0.565	0.415
France	0.700	0.435
Morocco	0.530	0.157
Netherlands	0.401	0.466
Portugal	0.583	0.320
Sweden	0.767	0.326
Tunisia	0.303	−0.030
Turkey	0.424	0.212

Note: Figures are Loevinger's H coefficients.

[7] Unfortunately, Mokken Scaling analysis cannot deal with missing values ('no opinion' responses), which we set at the scale midpoint. As these missing values are relatively uncommon, this should not affect our findings significantly.

should not be included in a spatial model because they do not have a sufficiently strong relationship with other items.

5. Scalability

The first aspect we examine is the extent to which the selected statements form a consistent scale. The coefficients of homogeneity of the left–right and progressive–conservative dimensions are displayed in Table 2. In all countries except the Netherlands the *H*-coefficient is higher for the left–right than for the progressive–conservative dimension. The average *H*-coefficient of the left–right dimensions is 'strong', while the average scalability of the progressive–conservative dimensions is only 'very poor'. In Canada, France and Sweden the left–right dimensions show 'very strong' *H*-values. The left–right dimensions in Egypt, Morocco and Portugal show high homogeneity coefficients. The left–right scale scores 'medium' in the Netherlands, Belgium and Turkey. Only in Tunisia the value falls below 0.4, which means that the scores are barely acceptable.

Using the same rules of thumb, we can assess the scalability of individual items. This allows us to see which items fit poorly in these scales. The left–right dimension in the Dutch *Kieskompas*, for example, includes four items that fit poorly. These concern the pension age, student loans and unemployment benefit. On these items, the traditional left–right divide is replaced by a division between those parties that support welfare state reform (including the left-wing GreenLeft and Labour parties) and those that oppose it (including the right-wing Freedom Party). The inclusion of these items in the left–right dimension reduces the overall scale homogeneity coefficient.

The *H*-coefficients of the progressive–conservative dimensions are lower than the coefficients from the left–right dimensions in most countries. The best scores are found in the Netherlands, France and Egypt - and even these do not exceed 0.5. In the other countries the scalability of the items included in the progressive–conservative dimensions is poor or very poor. The value for the Tunisian model is negative, which suggests that there is a negative association between many of the statements included in this dimension.

The reason that the progressive–conservative dimension is so weak in most countries is that these dimensions combine issues that do not tap into the same underlying dimension. One example is the Belgian statement on freedom of speech ('everybody should be allowed to say what he wants even if it is discriminatory'). The makers of this electoral compass assume that this is a progressive position. Empirically, this statement is supported most strongly by parties that are otherwise conservative, such as the Flemish Interest (*Vlaams Belang*). The theoretical idea behind the model which argues that freedom of speech is a progressive stance and the political reality in Belgium in which conservative parties support freedom of speech more contradict each other.

In the Arabic countries and Turkey, the progressive–conservative dimension concerns a range of issues including religious conservatism and foreign policy. The models assume that a conservative party is religious and takes an anti-Western stance on foreign policy. This relationship is, however, not empirically supported by the parties. In most cases there are anti-Western secular parties and religious parties that are centrist on foreign policy. This poses the question whether it is valid to combine these issues into a single dimension; at least from the 'supply side' perspective this is problematic.

All in all, the socio-economic left–right dimensions perform within acceptable parameters in almost all electoral compasses. The progressive–conservative dimension performs poorly or very poorly in most cases. The reason for this appears to be that different groups of items were taken together because they were assumed to tap into the same theoretical dimension, but they tapped into different latent empirical dimensions. In other cases, individual items were given a progressive or conservative direction for theoretical reasons, while empirically the relationships between these items and the rest of dimensions was opposite.

6. Dimensionality

Given that political parties may not always answer questions in the way that we expect theoretically, it may be useful to look at the models from an inductive perspective. While we do not argue that VAA models should be based solely on an inductive analysis of parties (or voter) positions on statements, it at least gives us an idea what kind of strong dimensions could be created. Moreover, it provides an opportunity to see whether 'flipping' items from their theoretically formulated direction may actually increase the scalability of our dimensions.

Our analysis aims to include all items in dimensions that form at least a poor scale ($H > 0.3$). In all countries except Canada and France, there are one or two 'unassigned' items. These relate so poorly to the other items that including them in any scale would reduce the scalability of that scale below acceptable levels. From a scaling perspective, they should have been excluded from the model.

The number of dimensions necessary to encompass all items differs from country to country (see Table 3). There is no 'one-size-fits-all' model that can be used in all studied cases. Although the Swedish *Valkompass*-model consisted of two dimensions, all but one item can actually be included in a very strong single scale. Still, not every item conformed to their theoretically formulated direction: two items were 'flipped' in direction. For example, the electoral compass model took a pro-introduction of the Euro position to be progressive, while empirically right-wing parties were more likely to support this. Thus in the inductive model, support of the introduction of the Euro is taken to be a measure of being right-wing rather than progressive. On the whole, this inductive analysis corroborates the idea that Swedish party politics is one-dimensional (Arter, 2008).

In a number of other countries the left–right dimension is by far the most important scale, which serves as a 'super issue' incorporating almost every political issue (Mair, 2007). In Canada, for example, 27 out of 30 items together form one very strong scale. The three remaining items can also be taken together to form a good scale as

Table 3
Assignment of statements to electoral compass and inductively created scales.

Electoral Compass	Inductive Dimension	Left–Right		Progressive–Conservative		*N*	*H*
		Items	Flipped	Items	Flipped		
Belgium	Dimension 1	13	0	14	2	27	0.571
	Dimension 2	5	0	1	0	6	0.506
	Dimension 3	0	0	2	0	2	0.724
	Unassigned	0	0	1	0	1	–
Canada	Dimension 1	12	0	15	2	27	0.731
	Dimension 2	1	0	2	0	3	0.656
Egypt	Dimension 1	14	1	4	1	18	0.585
	Dimension 2	0	0	10	0	10	0.828
	Unassigned	1	0	1	0	2	–
France	Dimension 1	14	0	13	1	27	0.653
	Dimension 2	0	0	3	0	3	0.772
Morocco	Dimension 1	10	0	16	4	26	0.464
	Dimension 2	2	0	1	0	3	0.412
	Unassigned	0	0	1	0	1	–
Netherlands	Dimension 1	11	0	11	1	22	0.629
	Dimension 2	6	0	1	0	7	0.498
	Unassigned	0	0	1	0	1	–
Portugal	Dimension 1	12	1	9	0	21	0.620
	Dimension 2	2	0	4	0	6	0.804
	Unassigned	1	0	0	0	1	–
Sweden	Dimension 1	14	0	15	2	27	0.714
	Unassigned	1	0	0	0	1	–
Tunisia	Dimension 1	8	1	12	4	20	0.426
	Dimension 2	1	0	3	1	4	0.488
	Dimension 3	1	0	3	2	4	0.484
	Unassigned	2	0	0	0	2	–
Turkey	Dimension 1	9	0	11	0	20	0.503
	Dimension 2	2	0	4	0	6	0.466
	Dimension 3	0	0	3	0	3	0.433
	Unassigned	1	0	0	0	1	–

Note: *N* indicates the number of statements included in the inductively created scale. *H* refers to Loevinger's coefficient of homogeneity.

well, although substantially these issues are very different: they concern the Senate, appointments to the Supreme Court and pensions.

In Egypt, Morocco, France, the Netherlands and Portugal two scales can capture most of the items, but these are not the same scales as used in the electoral compass model. In Morocco, most of the items from the social-economic left/right and the cultural progressive/conservative scale can be taken together into a single scale (although the direction of four of the items has to be changed compared to the original model). The resulting scale has a medium *H*-value. Out of the four remaining items, three can be included in a second dimension. Two of these concern policies towards the poor and unemployed and one concerns relations with Israel. In France, all but three items from a good scale. This scale includes economic and moral items. The three remaining items all concern EU integration and scale very well.

In the Netherlands and Portugal there are more items that do not fit into a single left–right dimension. Most items from the economic left/right and the cultural progressive/conservative dimension can be collapsed into one scale, with *H*-values that are higher than the left–right dimensions in the original specification. In Portugal, six items form a separate scale, these concern the European Union, the retirement age and the nationalization of banks. On these issues, the parties of the far left (the Left Block and Unitarian Democratic Coalition) differ from the centre-left Socialists and the parties of the (centre-)right. In the Netherlands, the inductively created left–right dimension consists of economic issues and cultural issues. In terms of party positions, immigration, integration and Islam do not form a separate dimension, but they confirm to the dominant left–right dimension (Otjes, 2011; Van der Brug and Van Spanje, 2009). Six items from the economic dimension actually do not cohere strong enough with the dominant left/right dimension to be integrated into this scale, including the welfare state reforms that we discussed above. Together they form a moderately strong scale. It divides reformist parties (especially the social-liberal D66) from the rightwing populist PVV and the leftwing populist SP, which tend to oppose welfare state reform.

Egypt conforms best to the electoral compass model. Fourteen out of fifteen items in the original left–right dimension form a consistent scale together with four items from the cultural dimension. Ten items from the cultural dimension form a consistent scale as well. This means that Egypt has a broad left–right dimension and a cultural dimension. The latter concerns Islam, religious tolerance, women's emancipation and censorship. This cultural

dimension separates the Muslim Brotherhood and other Islamic parties from the secular parties.

In the three remaining countries, Belgium, Tunisia and Turkey, three dimensions are necessary to incorporate (almost) all items. In Belgium, a broad left–right dimension can be constructed from a large majority of the items included in the model. Five social-economic items do not conform to the answering patterns in the social-economic dimension. Like in the Netherlands, most of these items concern welfare state reform. On these items the right-wing populist Flemish Interest sides with the left-wing social-democrats. Two of the three remaining items also form a very good scale as a pair: they concern defence and euthanasia. Turkey is in many ways a similar case: next to a moderately strong left–right dimension, which consists out of twenty of the thirty statements, there are two other dimensions consisting of a limited number of statements. These tap into a mix of different issues: the first includes foreign affairs (Cyprus and relations with other Islamic countries), forest management and civil liberties; the second also includes foreign affairs (relations with the United States), headscarves and constitutional reform.

The most problematic case from a scaling perspective is Tunisia. Most social-economic and most progressive/conservative items can be taken together into a single left/right scale (changing the direction of six items), resulting in a moderately strong scale. The remaining items can be assigned to two dimensions, which also scale quite low. It is difficult to interpret the substantive meaning of these scales.

7. The difference that empirically based spatial design makes

The last step in our analysis consists of comparing party positions in the original electoral compass models with our inductively obtained spaces. By comparing the two we show whether different modelling choices result in substantially different configurations of parties – and potentially different voting advice. Table 4 presents Pearson's coefficients of correlation between parties' positions on the two electoral compass scales (left–right and progressive–conservative) and their positions on our inductively created scales. We find, for example, that our inductively created first dimension in Belgium correlates strongly with the electoral compasses' left–right dimension and strongly negatively with the progressive–conservative dimension. This is not surprising given the fact that the correlation between the left–right and progressive–conservative dimension in the original electoral compass model is actually quite high. The first inductively obtained dimension reflects a broad left–right dimension incorporating socio-economic and cultural issues. The correlation between the second and third inductive dimension and the electoral compasses dimensions are lower. The same pattern can be found in Canada, Morocco, the Netherlands, Sweden and Turkey. In France and Portugal the correlations between the inductively created first dimension and the progressive–conservative scale is somewhat lower than in the aforementioned countries, but it is still quite substantial.

Table 4
Correlation between party positions on the electoral compass scales and the inductively created scales.

Electoral Compass	Inductive Dimension	Left–Right	Progressive–Conservative
Belgium	Dimension 1	0.85	−0.90
	Dimension 2	−0.79	0.18
	Dimension 3	0.32	−0.51
Canada	Dimension 1	0.99	−0.94
	Dimension 2	−0.18	0.42
Egypt	Dimension 1	0.97	0.14
	Dimension 2	0.14	−0.96
France	Dimension 1	−0.97	0.76
	Dimension 2	−0.33	−0.54
Morocco	Dimension 1	0.86	−0.87
	Dimension 2	−0.58	0.01
Netherlands	Dimension 1	−0.86	0.87
	Dimension 2	−0.52	−0.22
Portugal	Dimension 1	−0.96	0.74
	Dimension 2	0.59	0.25
Sweden	Dimension 1	0.98	−0.97
Tunisia	Dimension 1	0.78	0.19
	Dimension 2	−0.08	0.68
	Dimension 3	−0.51	0.10
Turkey	Dimension 1	−0.97	0.86
	Dimension 2	−0.27	−0.31
	Dimension 3	−0.24	0.52

Egypt is the only inductive model in which there are two substantive dimensions of which one taps into economic issues and the other taps into cultural issues. The first of these relates well to the electoral compass' (economic) left–right dimension, while the second correlates strongly with the progressive–conservative dimension. The off-diagonal correlations are very low. This confirms that in Egypt the electoral compass model fits best.

For Tunisia, the correlation between the inductive and deductive models is the lowest. Although the coefficients might seem quite high, this suggests that there are important differences between party positions on the inductively created scales and the original electoral compass scales. This further supports our argument that parties' answers to the Tunisian electoral compasses statements are difficult to model.

We should stress that we do not recommend that VAA creators solely rely on an inductive procedure to create their spatial models. The aim of our inductive analysis is to show that it is problematic to summarize parties' answers to VAA statements using the same two-dimensional model in each country. Our analysis suggests that in most countries one or two dimensions will be sufficient, but the content of these dimensions varies. In the Netherlands and Belgium welfare state reform may be an appropriate second dimension, while in Egypt it is religious conservatism. In other countries, such as the transitional political system in Tunisia, one may question the possibility to develop a coherent low-dimensional model altogether.

8. Conclusion

We argue that in addition to concerns about statement wording and selection, the underlying spatial model of VAAs should be based on scientific standards. VAA-makers should ensure that the statements that are used relate empirically to the dimension they are part of and do not tap into another substantive dimension. The models used in the electoral compasses in ten different countries are based on the designers' assumptions. Earlier studies have shown that these models to not conform to the voter space (Kleinnijenhuis and Krouwel, 2008). We demonstrated that party positions in many cases also fail to conform to these models: most of the models do not meet basic scalability requirements.

The good news is that the socio-economic left–right dimensions used in each of the national electoral compasses do form a moderate to strong scale, except in Tunisia. This does, however, not imply that it would be sufficient to *only* use the left–right scale as political research often does, at least not in VAAs. While left–right is clearly the most important dimension in most of the cases we studied, our analyses consistently show that in all countries except Sweden there is more going on than a single left–right dimension captures. Parties' positions on relevant political issues, such as secularism, European integration and welfare state reform would not be adequately reflected by using a single dimension.

Our problem mainly lies with the progressive–conservative dimensions that are used in the models: many of these have low to very low coefficients of homogeneity. If we inductively produce a low-dimensional model from parties' answers to the electoral compasses' statements, we find that the two-dimensional 'one-size-fits-all' model that was used in each of the countries does not fit well in all countries but Egypt. In the large majority of countries, we found a dominant left–right dimension comprised of both socio-economic as well as cultural issues. In all cases, except Sweden, one or two additional dimensions were found with very different interpretations (religious conservatism, welfare state reform, European integration). This suggests that the diversity of democratic party-systems cannot be captured in a uniform two-dimensional space.

This is problematic from the perspective of the goal of the electoral compass. It is supposed to inform citizens about the political space. Voters use it to understand the differences between political parties, especially between parties that are close to each other: voters consider multiple parties which are ideologically close and the vote compass helps them to make a decision between these parties (Wall et al., 2012). We argue that these models are only 'a' model of the political space based on one particular understanding of party competition and which, as we showed, only partially relate to an inductive party- or voter-based model. The models that are used do not reflect the differences between political parties correctly. This is not just a matter of scientific precision; these models play a role in democratic decision-making. Therefore they run the risk of misinforming the voters about which parties are closest to them. Earlier studies have found that the way in which the advice is calculated and specifically the choice of a spatial model has a very large effect on the vote advice that users receive (Kleinnijenhuis and Krouwel, 2008; Louwerse and Rosema, 2013). Creating appropriate spatial models is thus not (just) an academic exercise: it has a potentially large impact on the outcomes of VAAs.

Therefore, VAA designers should be transparent about all their modelling choices, including the reasons that they have for using specific dimensions and the number of dimensions. There might be compelling theoretical reasons to include statements into dimensions, which do not fit very well empirically (in terms of party or user answers). If so, these should be clear to the public.

References

Arter, D., 2008. Scandinavian Politics Today. Manchester University Press, Manchester.

Benoit, K., Laver, M., 2006. Party Politics in Modern Democracies. Routledge, London.

Benoit, K., Laver, M., 2012. The dimensionality of political space: epistemological and methodological considerations. Eur. Union Polit. 13 (2), 194–218 http://dx.doi.org/10.1177/1465116511434618.

Converse, P.E., 1964. The nature of belief systems in mass publics. In: Apter, D.E. (Ed.), Ideology and its Discontents. The Free Press of Glencoe, New York, pp. 206–263.

Coombs, C.H., 1964. A Theory of Data. Wiley, Hoboken.

Costa Lobo, M., Vink, M., Lisi, M., 2010. Mapping the political landscape: a vote advice application in Portugal. In: Cedroni, L., Garzia, D. (Eds.), Voting Advice Applications in Europe: the State of the Art. Scripta-Web, Napoli, pp. 143–185.

Cox, G., 2001. Introduction to the special issue. Polit. Anal. 9 (3), 189–191.

Downs, A., 1957. An Economic Theory of Democracy. Harper Collins, New York.

Fossen, T., Anderson, J., 2014. What's the point of voting advice applications? Competing perspectives on democracy and citizenship. Elect. Stud. 36, 244–251.

Hooghe, L., Marks, G., Wilson, C.J., 2002. Does left/right structure party positions on European integration? Comp. Polit. Stud. 35 (8), 965–989 http://dx.doi.org/10.1177/001041402236310.

Kleinnijenhuis, J., Krouwel, A., 2008. Simulation of decision rules for party advice websites. In: Paper Presented at the Annual Meeting of the American Political Science Association, Boston, Massachusetts.

Kleinnijenhuis, J., Scholten, O., De Ridder, J.A., Atteveldt, W., Van Hoof, A., Krouwel, A., Oegema, D., Ruigrok, N., Takens, J., 2007. Nederland vijfstromenland. De rol van de media en stemwijzers bij de verkiezingen van 2006. Bert Bakker, Amsterdam.

Krouwel, A., Fiers, S., 2008. From text to position. Text-extraction and authorisation of party positions along dimensions of political competition. In: Paper Presented at the Conference Voting Advice Applications (VAAs): Between Charlatanism and Political Science, Antwerp, Belgium.

Lefevere, J., Walgrave, S., 2014. A perfect match? The impact of statement selection on the voting advice application's ability to match voters and parties. Elect. Stud. 36, 252–262.

Louwerse, T., Otjes, S., 2012. Design challenges in cross-national VAAs: the case of the EU Profiler. Int. J. Electron. Gov. 5 (3/4), 279–297 http://dx.doi.org/10.1504/IJEG.2012.051305.

Louwerse, T., Rosema, M., 2013. The design effects of voting advice applications: comparing methods of calculating results. Acta Polit. http://dx.doi.org/10.1057/ap.2013.30.

Mair, P., 2007. Left–right orientations. In: Dalton, R.J., Klingemann, H.-D. (Eds.), The Oxford Handbook of Political Behaviour. Oxford University Press, Oxford, pp. 206–223.

Mokken, R.J., 1971. A Theory and Procedure of Scale Analysis. De Gruyter, Berlin.

Mokken, R.J., Lewis, C., 1982. A nonparametric approach to the analysis of dichotomous item responses. Appl. Psychol. Meas. 6 (4), 417–430 http://dx.doi.org/10.1177/014662168200600404.

Niemöller, K., Van Schuur, W., 1983. Stochastic models for unidimensional scaling: Mokken and Rasch. In: McKay, D., Schofield, N., Whiteley, P. (Eds.), Data Analysis and the Social Sciences. Frances Pinter Publisher, London.

Otjes, S.P., 2011. The Fortuyn effect revisited: how did the LPF affect the Dutch parliamentary party system. Acta Polit. 46 (4), 400–424 http://dx.doi.org/10.1057/ap.2011.12.

Van der Ark, L.A., 2007. Mokken scale analysis in R. J. Stat. Softw. 20 (11), 1–19.

Van der Brug, W., Van Spanje, J., 2009. Immigration, Europe and the 'new' cultural cleavage. Eur. J. Polit. Res. 48 (3), 309–334 http://dx.doi.org/10.1111/j.1475-6765.2009.00841.x.

Van Kersbergen, K., Krouwel, A., 2008. A double-edged sword! The Dutch centre-right and the 'foreigners' issue. J. Eur. Public Policy 15 (3), 398–414. http://dx.doi.org/10.1080/13501760701847655.

Van Schuur, W.H., 2003. Mokken scale analysis: between the Guttman scale and parametric item response theory. Polit. Anal. 11 (2), 139–163 http://dx.doi.org/10.1093/pan/mpg002.

Van Schuur, W.H., Kiers, H.A.L., 1994. Why factor analysis often is the incorrect model for analyzing bipolar concepts, and what model to use instead. Appl. Psychol. Meas. 18 (2), 97–110 http://dx.doi.org/10.1177/014662169401800201.

Wall, M., Krouwel, A., Vitiello, T., 2012. Do voters follow the recommendations of voter advice application websites? A study of the effects of Kieskompas.nl on its users' vote choices in the 2010 Dutch legislative elections. Party Polit. http://dx.doi.org/10.1177/1354068811436054.

Wheatley, J., 2012. Using VAAs to explore the dimensionality of the policy space: experiments from Brazil, Peru, Scotland and Cyprus. Int. J. Electron. Gov. 5 (3/4), 318–348 http://dx.doi.org/10.1504/IJEG.2012.051304.