## Original Article

# The design effects of voting advice applications: Comparing methods of calculating matches

Tom Louwerse[a] and Martin Rosema[b]

[a]Department of Political Science, Trinity College Dublin, 3 College Green, Dublin 2, Ireland.
[b]Centre for the Study of Democracy, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.

**Abstract**  In election times more and more voters consult voting advice applications (VAAs), which show them what party or candidate provides the best match. The potential impact of these tools on election outcomes is substantial and hence it is important to study the effects of their design. This article focuses on the method used to calculate the match between voters and parties. More specifically, we examine the use (explicit or implicit) of alternative spatial models and metrics. The analyses are based on the actual answers given by users of one of the most popular VAAs in Europe, *StemWijzer* in the Netherlands. The results indicate that the advice depends strongly on the spatial model adopted. A majority of the users of *StemWijzer* would have received another advice, if another spatial model had been used. At the aggregate level this means that how often a particular party is presented as best match depends strongly on the method used to determine the advice. These findings have important implications for the design of future VAAs.

## Introduction

In election times many voters turn to tests on the internet that tell them which party or candidate provides the best match, such as *smartvote* in Switzerland, *Wahl-O-Mat* in Germany and *StemWijzer* in the Netherlands. Especially in political systems where voters are faced with a multitude of parties or candidates such tools have become quite popular. In recent national elections in countries like Finland, Switzerland and the Netherlands, between 20 and 40 per cent of the electorate consulted a voting advice application (VAA) before casting their ballot (Garzia and Marschall, 2012). Typically, these tests ask individuals to indicate their personal opinions about a series of policy items, which are then compared with the positions taken by the competing parties or candidates. The outcome is a screen that presents the match or mismatch

between the individual and the alternative parties or candidates, which may be conceived of as advice to vote for the party with the best match.

The method that VAAs have adopted to determine the degree of match between voter and party or candidate differs (Cedroni and Garzia, 2010; Garzia and Marschall, 2012). One important element of the method is the type of spatial model that is used to calculate the match and to present the advice. For example, *Kieskompas* in the Netherlands plots the voter and parties in a two-dimensional political space. The implicit message is that the party closest in space provides the best match. The Swiss *smartvote* analyses policy preferences in terms of eight distinct dimensions and presents the results with a special diagram, which looks somewhat like a spider web. Even VAAs that do not present the results using a spatial framework, but for instance use bar charts with a separate column for each party, often adopt a particular spatial model for making the underlying calculations. We hypothesise that the method chosen, in particular the type of spatial model used, affects what party appears to be the best choice. Therefore, we examine whether, and to what extent, the choice for a particular spatial framework and metric influences the outcome of the tests.

The impact of the design of VAAs on their outcomes has been left largely unexplored. The research in this field, which is still not very extensive but quickly expanding, has focused primarily on the type of individuals that visit these websites (Wall *et al*, 2009; Fivaz and Nadig, 2010; Dumont and Kies, 2012), alternative methods to position political parties (Trechsel and Mair, 2011; Gemenis, 2013), links between theories of voting and VAAs (Mendez, 2012; Wagner and Ruusuvirta, 2012), whether the recommendations match pre-existing preferences (Wall *et al*, 2009), the type of parties that perform well in VAAs (Kleinnijenhuis *et al*, 2007; Ramonaitė, 2010), effects on electoral turnout (Fivaz and Nadig, 2010), and effects on party or candidate choice (Kleinnijenhuis *et al*, 2007; Walgrave *et al*, 2008; Dumont and Kies, 2012; Ladner *et al*, 2012; Wall *et al*, 2012). The only study that specifically addressed VAA design focused on the selection of statements (Walgrave *et al*, 2009). It found that statement selection has a profound impact on the match observed between party and voter. The selection of statements not only affected the outcome at the individual level for voters, but also at the aggregate level for parties: with certain sets of items particular parties turned out to provide the best match with many more voters than with other sets of items. The present article focuses on another aspect of the design, namely the method used to 'translate' answers to the statements into advice. We will demonstrate that this element of the design is at least as important, because it strongly influences what advice users get. The impact is also strong at the aggregate level: how well particular parties 'perform' in a VAA depends to a large extent on this design issue.

It is highly relevant to understand the effects of the methods employed by VAAs, because these tools may have a substantial impact on the outcome of elections. The last two decades in many western democracies VAAs have been introduced and

many voters take advantage of this opportunity and seek advice in the weeks or days, or sometimes even hours, before going to the polling station (Garzia and Marschall, 2012). The advice given to voters appears to have a significant impact on their vote choice, especially among floating voters (Walgrave *et al*, 2008; Ruusuvirta and Rosema, 2009; Ladner *et al*, 2010, 2012; Wall *et al*, 2012). This means that the potential impact of VAAs on the outcome of the election is substantial and hence it is important to study the effects of their design (Walgrave *et al*, 2009; Wagner and Ruusuvirta, 2012).

We examine to what extent the spatial framework adopted affects the advice given to voters. On the basis of an overview of methods employed by VAAs, we distinguish between four alternative spatial models. We determine how often the use of a different spatial framework would lead to different advice (another 'best match' between user and parties), as well as the degree to which parties match best with voters. The basis for our analysis are the actual answers given by millions of users of one of the most popular VAAs in Europe, *StemWijzer* in the Netherlands. In 2010 this test was completed almost 4.2 million times, which roughly corresponds with 40 per cent of the eligible voters.[1] The results to be presented below indicate that the spatial model matters a great deal. The choice for a particular spatial framework affects the outcome at two levels: for individual voters there are differences in terms of *which party* is presented as the best match, and for political parties there are differences in terms of *how often* they are presented as the best match to all voters. These findings have important implications for the interpretation of the results by voters and for the design of future VAAs.

## Voting Advice Applications and their Methods

One of the earliest VAAs is *StemWijzer*, which was developed for educational purposes by the Dutch non-profit organisation *Pro Demos*.[2] The first version of *StemWijzer* was created in 1989 as a paper-and-pencil test meant for high school education (De Graaf, 2010).[3] In 1998 the test became available on the internet and 6500 voters completed it online. In subsequent national elections the number of users grew rapidly: in 2002 the website was consulted 2.0 million times and in 2006 this number increased to 4.8 million. This amounts to about one-third of the total population and almost 40 per cent of those eligible to vote.[4] In 2010 the figure dropped to 4.2 million, probably because several other organisations developed their own online test. The most popular alternative is *Kieskompas*, which in 2006 was consulted about 1.5 million times (Kleinnijenhuis *et al*, 2007, p. 42; figures for 2010 have not been published, as far as we know). *StemWijzer* has also developed tests for elections at the municipal and provincial level, as well as for elections for the European Parliament. There are several spin-offs in other countries, such as *Wahl-O-Mat* in Germany and *Politarena* in Switzerland, which originate in

cooperation between the Dutch developers and non-profit organisations in those countries (Ladner *et al*, 2010; Marschall and Schmidt, 2010).

The design of *StemWijzer* has undergone some changes throughout the years. The first version contained 60 ideological statements derived from election manifestos (De Graaf, 2010). In later tests the number of statements was reduced to 30, which were taken from a long list of about 50 statements for which party positions would be determined. The basis for positioning the political parties is election manifestos and consultations with the party headquarters. When compiling the final list of items, important criteria are that the statements involve some controversy (there should be parties in favour as well as against and most parties should take a position) and that the list as a whole provides sufficient opportunity to distinguish between any pair of parties. For each statement parties as well as users have three alternative positions: agree, neutral (or 'neither') and disagree. In addition, users have the possibility to indicate no opinion and skip the item.

In the first paper-and-pencil test the statements were used to plot users and parties on a left/right continuum. In later years, however, the use of this simple spatial model was abandoned. In all online versions of *StemWijzer* the match between party and user has been determined without the explicit use of a spatial model. Between 1998 and 2006, for each party the mismatch was calculated by awarding 2 points if party and user take opposite stands on a statement (agree versus disagree) and 1 point if either of them is neutral whereas the other agrees or disagrees (De Graaf, 2010). Furthermore, users could identify any number of statements that they consider more important and these items are then given extra weight (points are doubled). The match with each party is calculated as a percentage of agreement (percentage of maximum points subtracted from 100 per cent). On the final screen the user is informed which party provides the best match and parties are ranked according to the degree of match, which is visualised with bar charts (see Figure 1(a)).

A key characteristic of the method employed by *StemWijzer* is that it does not explicitly use a spatial framework for making calculations or presenting the advice (except for the original paper-and-pencil test). However, the method may still be conceived of in those terms. It basically corresponds with a spatial model in which each item represents a separate dimension. The metric used is city block distance, which deviates from the Euclidean distance measure that is more common in spatial models.[5] Since 2006, *StemWijzer* has used an even simpler 'agreement method': one point is awarded if the position of party and user are identical, while no points are awarded if both have different positions. For example, if a user is neutral and a party disagrees with a statement, no points are awarded. Double points are awarded for issues that have been given extra weight by the user. The best match is the party with most points.

The method used by *StemWijzer* is straightforward and relatively easy to understand, which is an important characteristic from the perspective of users. However, inspired by research in political science in which political parties are mapped in
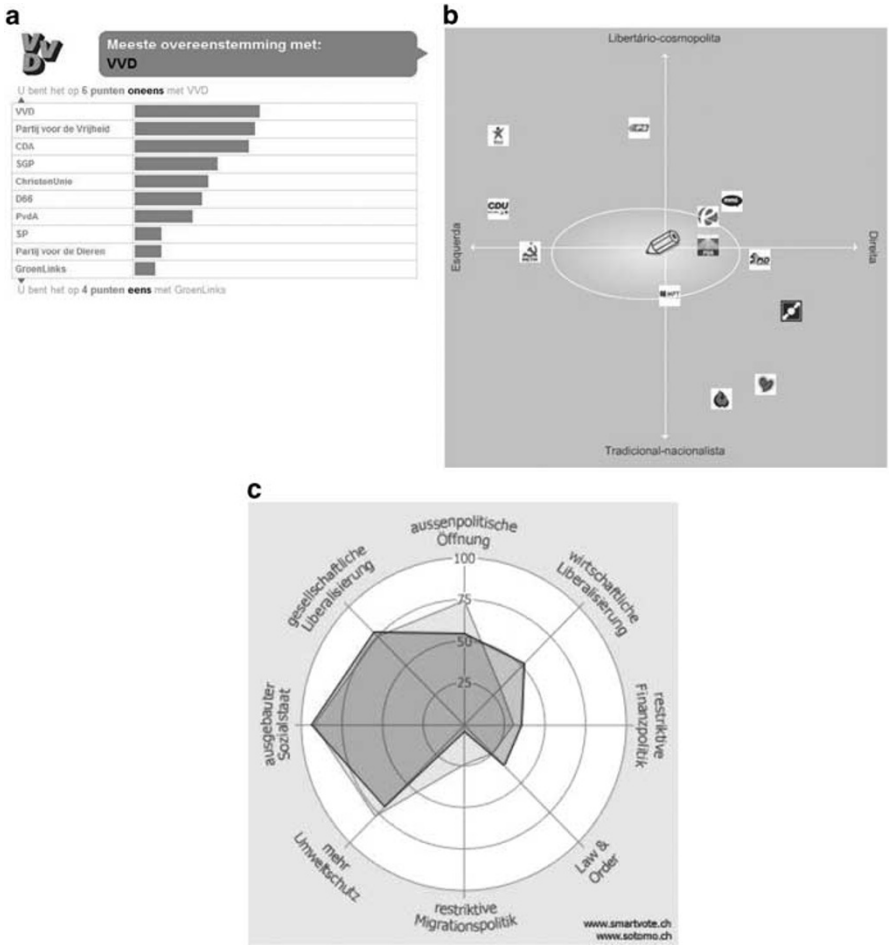
**Figure 1:** Result screens of voting advice applications (VAAs): **(a)** Dutch VAA *StemWijzer*; **(b)** Portuguese VAA *Bússola Eleitoral*; **(c)** Swiss VAA *smartvote*.

a so-called political space, many VAAs adopt a different and somewhat more complicated approach: they use the items to create a spatial model in which voters and parties are positioned (cf. Downs, 1957; Enelow and Hinich, 1984). The idea behind the spatial representation is that it provides voters with insight in the policy positions of political parties and the differences between parties, rather than 'just' matching user preferences with those of the parties. Moreover, it may stimulate voters to consider the coherence of policy preferences, that is, that it is inconsistent to support both lower taxes and more government spending. The implicit advice of

these VAAs seems to be to cast a vote for the party that is closest to oneself in the political space, analogous with the smallest distance hypothesis in spatial models of voting (see Downs, 1957). In the Netherlands this procedure has been adopted by *Kieskompas*. Its methodology has subsequently been applied in many other countries and was also used, among other methods, in *EU Profiler* (Trechsel and Mair, 2011, Figure 5).

An example of a VAA that was modelled in this way, and for which the method has been documented in some detail (Lobo *et al*, 2010), is the Portuguese *Bússola Eleitoral* ('electoral compass'). In the run-up to the 2009 national elections this test was completed about 175 000 times, which corresponds with 3 per cent of the electorate. The developers identified a socio-economic dimension of left/right as the major basis for party competition in Portugal. In addition, they formulated a second dimension that resembles the division between Green/Alternative/Libertarian (GAL) and Traditional/Authority/Nationalist (TAN) (Marks *et al*, 2006). In total, 28 statements were selected, of which 15 linked up with left/right and 13 linked up with GAL/TAN. The answer format was a 5-point rating scale ranging from 'completely agree' to 'completely disagree'. The positions of users and parties on each dimension were determined by taking the average position on each item related to that dimension, using a scale from –2 to +2. Party positions were determined on the basis of election manifestos, debates in parliament, and statements on party websites and in the media. The website also asked users to rate party leaders and indicate the chance they would ever vote for each party, but answers to these questions were not used for calculating the match.

The match between user and party was determined by *Bússola Eleitoral* in two ways. First, on the basis of answers to the statements the position of users and parties on both axes of the political space were determined and plotted on the screen (see Figure 1(b)). The implicit message is that the party closest in this space (Euclidean distance) provides the best match. Second, agreement scores were determined by calculating the average agreement on all 28 statements.[6] Lobo *et al* (2010, p. 168) argue this is actually a better measure: 'These scores are more accurate because, hypothetically, it could well be the case that a user occupies the same position on both of the axes, as one of the parties, but that this proximity is based on agreement and disagreement with different statements from the questionnaire'. This is indeed true, although the question remains to what extent in practice both methods lead to different results.

The most popular VAA in Switzerland, *smartvote*, presents the results in three different ways. Next to the two methods presented above, it uses the so-called *smartspider*: a model that combines eight policy dimensions (see Fivaz and Schwarz, 2007; Ladner *et al*, 2010). *Smartvote* appeared on the internet in 2003, when it was consulted 250 000 times. In 2007 the usage figure increased to 1.0 million, which corresponds with about 40 per cent of the Swiss electorate. The spider diagram of *smartvote* comprises eight axes with values ranging between 0 and 100.[7] Each voter

and party (or candidate) is positioned on each dimension. The points that represent the positions of a voter are connected by lines, which results in an area in the diagram. The same is done for parties (or candidates, in the case of *smartvote*). By comparing the areas of voter and party, which can be done in the same diagram, one can see the amount of overlap and hence the degree of match between both (see Figure 1(c)). If the areas overlap completely, voter and party take identical positions on each dimension and the match is perfect. The less overlap there is between the voter area and the party area, the poorer the match between both. To find out which party provides the best match the graphs for all parties (combined with the voter) have to be compared by the user.

The above overview is not an exhaustive list of all methods employed by VAAs, but we believe it captures the essential differences between VAAs with respect to the methods used to determine the match between voter and party and present the advice. Put simply, there are four methods. The first method, which is central in *StemWijzer*, corresponds with counting the proportion of statements where voter and party agree. The best match is provided by the party with whom the voter agrees most often. This method implicitly adopts a multidimensional spatial framework in which each item represents a separate dimension. This method can be applied with different metrics: city block distance, Euclidean distance or the agreement method used in the most recent editions of *StemWijzer*. The second method is the one that *StemWijzer* started with in 1989, namely by constructing a one-dimensional space, like a left/right continuum, and position voters as well as parties on this scale. The third method, which has been used by *Kieskompas* and its spin-offs, consists of positioning voters and parties in a two-dimensional space. These positions are typically determined by averaging the scores on the items that are considered as indicators of these policy dimensions. The fourth method also adopts a spatial framework, but distinguishes multiple dimensions. *Smartvote*, which identifies eight policy dimensions, is a typical example.

In this article we compare these spatial models and examine if the use of different models would lead to different advice at the individual level as well as at the aggregate level. In addition, we analyse if the use of a different metric (Euclidean distance, city block distance, agreement method) affects the result. In the next section we describe the data and method in some detail before and then proceed with the analysis.

## Data and Method

The analysis we present is based on the responses given by the 4.2 million users of the *StemWijzer* edition for the Dutch parliamentary election in 2010. The Netherlands is a suitable case for analysing the method of calculating matches, because of the widespread use of VAAs and the multi-party system. If there are only two parties, the

method will likely only have limited effect, but with a multitude of parties the effects are presumably larger. Moreover, countries with a multi-party system have traditionally been frontrunners in VAA use, exactly because voters need to weigh so many alternatives. Therefore, it is most crucial to observe design effects in such cases. *StemWijzer* is the most widely used VAA in the Netherlands and thus provides a rich data set to explore this issue.

The 30 items included in the 2010 edition of *StemWijzer* are listed in Appendix A. The log files of the online application contain information on the positions taken on each statement as well as the extra weight allocated to each statement by users, as well as the party positions. This allowed us to compare the results of different methods to calculate the match between voters and parties.

Answer profiles that only contained missing answers ('skip this question') were excluded. In addition, we excluded about 30 000 recommendations because we suspect that these were computer-generated. These represent three cases where one single IP address requested advice thousands of times giving exactly the same answers to the statements. Next, we took a random sample of 10 000 cases for further analysis, because analysing the full data set would be too memory-intensive. Although this may introduce some error, a sample of this size will yield results that are almost certainly extremely close to what would have been found using the full data set. Most users in our sample (90 per cent) answered all 30 statements and very few skipped more than five questions (1 per cent). A majority (77 per cent) made use of the opportunity to select statements that they considered particularly important (mean = 5.1; standard deviation = 4.7).

The 2010 edition of *StemWijzer* by default included the 11 political parties that were represented in the Second Chamber of the Dutch parliament at that time. Voters had the opportunity to deselect any of these parties, while they could also select any of six additional parties for which the application included data. The log files do not contain information on which parties were (de)selected by each user, but the low number of recommendations for parties that were not included by default (2 per cent) suggests that only few people used this possibility. Because we are interested in the effect of VAA aggregation procedures on the voting advice for users, we opted to include the default selection of parties in our analysis. This presumably most truthfully reflects the advice a majority of users received.

## Alternative methods of calculating voting advice

We implemented eight methods of calculating voting advice on the basis of the spatial models and metrics discussed in the preceding paragraphs: (a) high-dimensional agreement method (the method that *StemWijzer* used in its 2010 edition), (b) high-dimensional city block distance method, (c) high-dimensional Euclidean distance method, (d) one-dimensional model, (e) two-dimensional model,

(f) three-dimensional model induced from parties' answers to the statements, (g) three-dimensional model induced from users' answers to the statements, and (h) seven-dimensional 'spider' model. The implementation of the first three methods is relatively simple (see Appendix B). One calculates the agreement or distance between the answers of a particular user and those of each party, ignoring missing answers on the part of the user. Extra weights allocated by users increase the agreement or distance between users and party. For the agreement method this results in an agreement score, which is highest for the 'best match'. The city block and Euclidean distance method result in (weighted) distance scores, which are lowest for the 'best match'. We included the extra weights voters could put on statements in these models (as this has been standard practice in StemWijzer and other VAAs), but our findings would have been similar if these weights would not have been taken into account.

The one-, two-, and multidimensional models require a method of combining items into issue dimensions. For the one-dimensional model this is a matter of determining the direction of each statement: does agreeing imply a left-wing/progressive position or a right-wing/conservative position? We determined this on *a priori* grounds and checked the homogeneity of the resulting policy dimension using Loevinger's $H$. When looking at the answers of the political parties, this resulted in a scale with $H = 0.37$, which is low but acceptable. For the users, the homogeneity coefficient is very low ($H = 0.07$), which indicates that for voters a one-dimensional approach is insufficient. Nonetheless, we include this method to demonstrate what its adaptation would mean. Moreover, some scholars have argued that a one-dimensional left-right model is suitable in the context of electoral choice (Downs, 1957; Van der Eijk and Niemöller, 1983).

The two-dimensional model closely follows the method of *Kieskompas*. The model consists of a socio-economic left-right dimension and a progressive-conservative (GAL/TAN) dimension (Marks *et al*, 2006) and each statement has been assigned to either dimension (see Appendix A). The resulting scales are not very strong with $H$ coefficients of 0.20 and 0.40 respectively for the parties' answers, and very weak $H$ coefficients of 0.06 and 0.07 for the user data. Similar results concerning the strength of the Kieskompas model have been found by Otjes and Louwerse (2011, pp. 10–13), who analysed the items in this VAA.

The two-dimensional model has been constructed by selecting the relevant policy dimensions *a priori*. However, one could argue that first relevant political issues should be selected and that the appropriate spatial model should be induced from the patterns of (party or voter) answers given to these statements (Otjes and Louwerse, 2011). We may thus find out inductively that parties' answers to the statements can be captured well by a one-dimensional or two-dimensional spatial model. This type of model was fitted using classical multidimensional scaling (MDS). This method uses a Euclidean distance measure between actors (parties or users), based on their answers to the VAA statements, and tries to find a low-dimensional approximation of

those distances. We applied this method in two ways, namely once on the basis of party positions and once on the basis of voter positions. The degree to which a low-dimensional model accurately represents the distances between parties is measured by Kruskal's Stress-I statistic. Stress levels below 10 per cent are considered acceptable. For the data set of party responses to the statements a three-dimensional solution was found to be acceptable (Stress = 4.37). The next step was to determine to which dimension each statement was connected, which was determined by regressing parties' answers to each statement on these three dimensions, a technique called property fitting. We included an item in the dimension that provided the highest beta coefficient for that particular item, provided the $R^2$ was larger than 0.30.[8] In this way 29 out of the 30 items were included in one of the scales (see Appendix A). We use additive scales to construct the model, because this reflects most closely how other VAAs, such as *Bússola Eleitoral* and *Kieskompas*, construct their spatial model. An additional advantage of this method is that for users it is more transparent than more sophisticated techniques such as factor analysis or MDS. The resulting scales had high $H$ values of 0.71, 0.75 and 0.54, respectively (based on parties' answers). However, when applied to the users' answers these scales are not very strong ($H = 0.09$, 0.06 and 0.15).

In an alternative specification users' answers to the statements were used in an MDS analysis. Thus, instead of a 'party space', a 'voter space' was constructed. Answer patterns of users proved to be more erratic than those of parties: a three-dimensional solution had a stress level of 30 per cent, but including more dimensions only reduced this level very gradually to 13 per cent for a 10-dimensional model. For reasons of clarity and comparability, we decided to stick to a three-dimensional model in this case as well, despite the poor fit. After all, the logic behind a low-dimensional model of party positions is to provide users with insight in the different policy stances of parties – presenting a 10-dimensional model would destroy this objective. Three dimensions have also been induced by Aarts and Thomassen (2008) based on their analysis of voters' evaluation scores of political parties. The H coefficients for the three dimensions were 0.29, 0.08 and 0.13, respectively, which is (somewhat) better than the $H$ values (for voters) obtained from the party space, as one might expect. Still, the scalability of these items is low. To determine the resulting voting advice from the multidimensional models, the distance between users and parties were calculated as (unweighted) Euclidean distances.

The last method, a seven-dimensional model that reflects the spider diagram, has been implemented by assigning issues to one of the seven categories that were used in the *EU Profiler's* spider diagram (see Appendix A; Trechsel and Mair, 2011, Figure 6). The assignment was based on *a priori* grounds and checked using the homogeneity coefficient *H*. For the answers provided by the parties, the *H* values for each of the seven issue dimensions were over 0.3, except for welfare state politics. Although the coefficient for this category could be improved by changing the direction of some of the items, this would run contrary to the substantive meaning of

the category. The *a priori* approach seems to fit most closely to how *smartvote* and *EU Profiler* construct their spider models (note that *smartvote* uses eight categories, but the principle is the same). Users' and parties' answers to the statements were recoded and summed up, so that a score of 100 on a particular axis indicates complete agreement and 0 indicates complete disagreement. The total distance between voters and parties was calculated as the city block distance, which seems to fit most closely to the way a spider diagram represents the voting advice.

## Three measures to compare the models

We use three different measures to compare the advice stemming from the alternative models. The first measure focuses on the party that provides the best match and indicates how often two given methods provided the same 'best match'. If an individual received the advice 'Freedom Party' (PVV) using the city block metric as well as the Euclidean metric, this constitutes a full match. There are also cases where the 'best match' was a tie between two or more parties. When at least one party was among the best matches in both methods, this was regarded as a partial match. All other cases were treated as 'no match'. Although a benchmark for this measure is somewhat arbitrary, one could argue that we should be able to observe at least two matches versus one mismatch. This would correspond to 67 per cent (full) matches.

The second measure focuses on the degree of match between the user and each individual party, thus looking broader than only which party provided the best match. To capture the similarity of the advice, we calculated a correlation coefficient between the match scores of two methods for each individual party. For example, if there is a perfect linear relationship between the Labour Party scores according to the agreement method and the city block method, the correlation coefficient equals one. To estimate the overall similarity between two methods we take the means of these correlations across the 11 parties. These average correlation coefficients should be rather high, given the fact that the various methods are all based on the same data and after the same outcome. Correlation coefficients of 0.7 or higher, which corresponds to roughly 50 per cent explained variance or more, should be achievable.

The third measure concerns the number of times that each party was recommended at the aggregate level (where tied recommendations are divided between the parties concerned). Some methods may divide the recommendations more evenly across parties, while other methods may favour specific parties. Furthermore, it is possible that specific parties 'benefit' from a particular method. In contrast to other studies (Kleinnijenhuis and Krouwel, 2008; Walgrave *et al*, 2009) we do not take the election result as the 'gold standard' for comparison. After all, the aim of a VAA is not to predict or mimic the election result, but to inform voters about their substantive policy match with parties. Voters may well decide to vote on other grounds. Nevertheless, a comparison between the number of recommendations and the actual

number of votes may be considered interesting, because it indicates to what extent the electorate supports parties that most closely represent their views on a wide range of policy issues. More importantly, if we find large differences between the proportion of recommendations for particular parties across the alternative methods, this provides clear evidence for our hypothesis that the method used to calculate the advice matters.

## Results

### Effects at the individual level

The eye-catching element of any VAA result page is which party provides the 'best match' for a user. Table 1 displays the similarity between the 'best matches' that users would have received under different methods of calculating the result. The similarity between the agreement, city block and Euclidean methods is high: for each dyad, over 80 per cent of users would receive exactly the same 'best match'. Given the fact that each of these methods treats the statements as independent and adds up differences between parties and voters on each statement, this is in line with our expectation. Nonetheless, the advice is not exactly the same. Because the 'best match' of one method is in many cases only slightly better than the 'second best match', small differences in the calculation method may affect which party appears on top of the list.

The agreement method does not always provide the user with a single 'best match'. In about 20 per cent of the cases the 'best match' was a tie between two or more parties. The percentage of users with tied recommendations was lower for the city block and Euclidean methods, 12 and 9 per cent respectively, and almost non-existent for any of the spatial methods (1 to 3 per cent). The ties influenced the degree of similarity between advice based on agreement scores, city block metric and Euclidean distance: in most cases where no full match was observed, a partial match existed (for 99 per cent of the users a partial match between agreement and city block method was observed, while for agreement method and Euclidean distance the corresponding figure was 94 per cent). Because ties were uncommon in the one-dimensional and multi-dimensional spatial models, the comparisons between those are not strongly affected by the presence of ties.

The similarity between the various spatial models and the three high-dimensional approaches is low. This is certainly true for the one-dimensional model: only 6 per cent of users would receive the same recommendation on the basis of the one-dimensional model and high-dimensional agreement method (and for only 5 per cent there is a partial match). This result alone already shows that it greatly matters which underlying model is used. The other spatial approaches show a somewhat higher similarity with the agreement method, but also reveal strong differences.

**Table 1:** Percentage of users who received the same 'best match' by two methods

| | Agreement | City block | Euclidean | One-dimensional | Two-dimensional | Three-dimensional (Party MDS) | Three-dimensional (Voters MDS) | Seven-dimensional ('Spider') |
|---|---|---|---|---|---|---|---|---|
| Agreement | 100.0 | 85.2 | 83.0 | 6.4 | 23.2 | 13.0 | 35.2 | 41.3 |
| City block | 85.2 | 100.0 | 94.7 | 7.7 | 24.6 | 16.0 | 37.2 | 43.7 |
| Euclidean | 83.0 | 94.7 | 100.0 | 9.1 | 25.0 | 16.7 | 37.6 | 43.9 |
| One-dimensional | 6.4 | 7.7 | 9.1 | 100.0 | 19.9 | 18.4 | 12.4 | 9.5 |
| Two-dimensional | 23.2 | 24.6 | 25.0 | 19.9 | 100.0 | 22.2 | 29.8 | 26.9 |
| Three-dimensional (Party MDS) | 13.0 | 16.0 | 16.7 | 18.4 | 22.2 | 100.0 | 17.2 | 22.1 |
| Three-dimensional (Voters MDS) | 35.2 | 37.2 | 37.6 | 12.4 | 29.8 | 17.2 | 100.0 | 38.4 |
| Seven-dimensional ('Spider') | 41.3 | 43.7 | 43.9 | 9.5 | 26.9 | 22.1 | 38.4 | 100.0 |

*Abbreviation*: MDS, multidimensional scaling.

The two-dimensional model provides the same best match as the agreement method in 23 per cent of the cases (with 9 per cent partially matching). The three-dimensional space that was constructed on the basis of parties' positions gives the same advice as the agreement method in only 13 per cent of the cases (while for 9 per cent it gives a partial match). The three-dimensional voter model does the same for 35 per cent of the cases. The seven-dimensional 'spider' model is somewhat more congruent with the agreement method: 41 per cent of users receive a similar best match, while another 13 per cent has a partial match between both methods. Thus, as the dimensionality of the underlying model goes down, the advice gets more different from the agreement method.

Apart from the difference between high-dimensional and low-dimensional models, there is also a difference within the group of low-dimensional (spatial) models. For example, the 'best matches' provided by the three-dimensional model based on parties' positions is only in 19 per cent of the cases the same as the advice provided by the three-dimensional model based on voters' positions. It matters a great deal how a model is constructed.

While the 'best match' is informative, it only takes into account one particular aspect: the top of the stack. It is possible that shifts in advice stem from rather small differences in the scores of parties. Therefore, we have calculated how similar the results are under different methods by looking at the correlation between party scores for any two methods (see Table 2). The figures display the pattern that was already observed in Table 1. The agreement, city block and Euclidean method show largely similar advice, while the average correlation between the one-dimensional and other methods of calculating the voting advice is moderate. For example, the correlation between the scores based on the agreement method and the three-dimensional spatial model amounts to about 0.6, which is below the minimum benchmark level of 0.7 that we set above. This means that the amount of explained variance amounts to less than 40 per cent. So the main observation is that the correlations are rather low, given the fact that the different methods all aim at the same outcome (determining how well the policy positions of a party matches with opinions of a voter) using the same data. These figures underline that the method used to calculate matches strongly affects to what extent users and parties are perceived as holding similar policy positions.

## Effects at the aggregate level

The final question is whether the differences in advice at the individual level translate into differences at the aggregate level. Does the number of recommendations to vote for a particular party differ between methods, or do the differences at the individual level cancel each other out? Table 3 shows this, while also presenting the actual election outcome.

**Acta Politica** Vol. 49, 3, 286–312

**Table 2:** Mean correlation between the matching scores of two methods (Pearson's *R*)

|  | Agreement | City block | Euclidean | One-dimensional | Two-dimensional | Three-dimensional (Party MDS) | Three-dimensional (Voters MDS) | Seven-dimensional ('Spider') |
|---|---|---|---|---|---|---|---|---|
| Agreement | 1.00 | 0.91 | 0.79 | 0.51 | 0.51 | 0.58 | 0.57 | 0.60 |
| City block | 0.91 | 1.00 | 0.97 | 0.58 | 0.58 | 0.66 | 0.66 | 0.69 |
| Euclidean | 0.79 | 0.97 | 1.00 | 0.56 | 0.56 | 0.64 | 0.66 | 0.69 |
| One-dimensional | 0.51 | 0.58 | 0.56 | 1.00 | 0.80 | 0.87 | 0.68 | 0.60 |
| Two-dimensional | 0.51 | 0.58 | 0.56 | 0.80 | 1.00 | 0.75 | 0.62 | 0.62 |
| Three-dimensional (Party MDS) | 0.58 | 0.66 | 0.64 | 0.87 | 0.75 | 1.00 | 0.67 | 0.64 |
| Three-dimensional (Voters MDS) | 0.57 | 0.66 | 0.66 | 0.68 | 0.62 | 0.67 | 1.00 | 0.67 |
| Seven-dimensional ('Spider') | 0.60 | 0.69 | 0.69 | 0.60 | 0.62 | 0.64 | 0.67 | 1.00 |

*Abbreviation*: MDS, multidimensional scaling.
*Note*: These are the means of correlations for each party.

**Table 3:** Percentage of best matches for each party, by method

| | Agreement | City block | Euclidean | One-dimensional | Two-dimensional | Three-dimensional (Party MDS) | Three-dimensional (Voters MDS) | Seven-dimensional ('Spider') | Election result 2010 |
|---|---|---|---|---|---|---|---|---|---|
| Liberal Party (VVD) | 6.8 | 7.0 | 6.7 | 1.1 | 0.7 | 0.6 | 2.9 | 1.4 | 20.5 |
| Labour Party (PvdA) | 18.8 | 17.7 | 17.0 | 1.2 | 2.9 | 0.8 | 13.2 | 17.5 | 19.6 |
| Freedom Party (PVV) | 34.0 | 32.8 | 31.8 | 3.5 | 36.6 | 0.6 | 43.3 | 32.5 | 15.5 |
| Christian Democrats (CDA) | 4.8 | 3.9 | 3.4 | 8.8 | 1.1 | 0.0 | 1.2 | 0.4 | 13.6 |
| Socialist Party (SP) | 9.8 | 9.7 | 9.1 | 1.2 | 6.3 | 20.3 | 8.4 | 3.5 | 9.8 |
| Democrats 66 (D66) | 3.5 | 4.5 | 5.3 | 1.2 | 12.7 | 1.5 | 3.9 | 2.2 | 7.0 |
| Green Left | 3.6 | 3.1 | 2.9 | 0.5 | 0.4 | 0.5 | 0.7 | 1.9 | 6.7 |
| Christian Union | 3.4 | 3.0 | 2.7 | 27.2 | 27.8 | 27.0 | 5.1 | 3.7 | 3.2 |
| Dutch Reformed Party (SGP) | 2.7 | 4.1 | 6.0 | 52.7 | 5.0 | 3.9 | 9.2 | 6.1 | 1.7 |
| Party for the Animals | 5.1 | 6.2 | 6.7 | 0.3 | 0.5 | 3.4 | 3.5 | 4.4 | 1.3 |
| Proud of the Netherlands (TON) | 7.9 | 8.8 | 9.0 | 2.7 | 6.4 | 42.6 | 9.1 | 26.8 | 0.6 |

*Abbreviation*: MDS, multidimensional scaling.

*Note*: In cases where two or more parties were tied in first place, the advice was divided evenly over the parties concerned.

The agreement method of *StemWijzer* produced 34 per cent of the recommendations for the Freedom Party (PVV) led by Geert Wilders, which thus was the party that most often matched best with users' policy preferences. The figure for the Labour Party (PvdA) was 19 per cent, while all other parties score below 10 per cent. In a sense the high figure for Wilders' party is not surprising, because research has shown that many voters combine left-wing policy preferences on the socio-economic left-right dimension with right-wing policy preferences on the cultural left-right dimension (Van der Brug and Van Spanje, 2009). The Freedom Party approximates this position most closely. The aggregate outcome for the 11 parties would not have been very different would *StemWijzer* have used the city block metric or Euclidean distances: the percentages for each party are nearly identical. One exception is the small orthodox Protestant party SGP, which received 2.8 per cent of best matches in our sample according to the agreement method, but would have more than doubled that with the Euclidean distance method. Yet, on the whole the differences are limited.

If we compare the high-dimensional agreement method with the other spatial models, sizeable differences can be observed. The one-dimensional model provides the most deviant outcome. With this method two small parties would have benefitted: 53 per cent of best matches would concern the SGP and 27 per cent the other small Protestant party Christian Union. This remarkable outcome can be explained by looking at the distribution of parties and voters on this single dimension, which is displayed in Figure 2. The top half of the figure displays the distribution of the users in our sample by means of a violin plot, a combination of a box-plot displaying the mean, quartiles and (truncated) range of values, and a density plot. It shows that 50 per cent of the users scored between −0.167 and +0.167 on the continuum, which ranges from −1 to +1. The parties' positions are located much more towards the extremes and in two clusters: one left-wing and one right-wing. Thus, a large majority of users has a left-right position somewhere in between the left block and the right block. As a result, the parties in either block that are located closest to the centre receive this advice. For the left-wing block this is the Christian Union and for the right-wing block the SGP. The paradoxical result is that the parties which fit the left-right framework most poorly, the two orthodox Protestant parties (Pellikaan *et al*, 2003), would benefit most strongly from adapting this method for calculating the VAA result.

The two- and three-dimensional models suffer largely from the same problem as the one-dimensional model: voters are clustered in the centre and therefore parties which are close to the centre do well in terms of 'best matches'. Which parties are close to the centre depend strongly on the model adapted. In the two-dimensional model, Freedom Party (PVV), Christian Union (CU) and Democrats 66 (D66) are close to the centre; for the Party MDS model it is the Socialist Party (SP), Christian Union (CU) and Proud of the Netherlands (TON), while for the Voter MDS model, the Freedom Party (PVV) and Labour (PvdA) are located more towards the centre. This is reflected in the number of best matches for these parties. The seven-dimensional 'spider' representation also suffers from this problem, but to a more
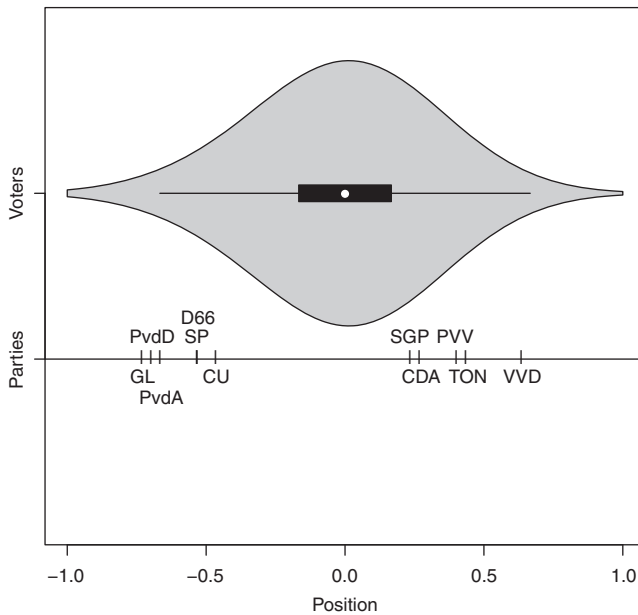
**Figure 2:** Parties' and users' positions in the one-dimensional model.
*Note*: Voters' positions are based on a sample of 10 000 VAA users. The voters' plot is a violin plot, which combines a boxplot and a density plot.

limited extent. Compared with the agreement method, one party scores much better in this model, Trots op Nederland (TON), while most of the other parties do slightly worse.

It is clear that the election result does not accurately reflect the recommendations of *StemWijzer*, or vice versa. This is well illustrated by the fact that the election winner, Liberal Party (VVD), received only 7 per cent of the recommendations. The figures indicate that with another metric or another spatial model the situation would have been the same. None of the methods would have resulted in an aggregate level outcome that matches the actual election result. This is not at all worrisome, however, because the aim of VAAs is not to predict election outcomes and voters may have many other reasons to support a particular party (for example, social influence, habit, religious background, government approval, party leader effects, and so on).

## Conclusions

VAAs have become an important tool for voters, enabling them to find out how well their policy preferences match with the positions taken by political parties or candidates. Given their widespread use, the potential impact on election outcomes is substantial. Against this background we have examined an important element of the

design of VAAs, namely the spatial framework that VAAs adopt to translate user answers into advice. Previous research by Walgrave *et al* (2009) demonstrated that statement selection by VAAs has a substantial impact on the nature of the advice. We have shown that with the same set of statements, the method to calculate the match between voter and party also has a strong influence on the outcome of the test.

The analyses presented are based on the actual answers by users of the 2010 edition of *StemWijzer*, a Dutch VAA that was consulted by about one-third of the voters in the run-up to the national election. We compared the 'agreement method' used by the original application, which corresponds with a simple count of the proportion of statements where user and party take the same position, with several spatial models that are central in other VAAs. In addition, we analysed the impact of the type of metric used to calculate the match: Euclidean distance, City block distance and *StemWijzer*'s agreement method. The analyses show that the design matters a great deal. Up to 90 per cent of users would have received a different voting recommendation if another method for calculating the match between voter and party had been adopted. Even for the spatial model that gave the most similar results as the original application – that is, a seven-dimensional model inspired by the Swiss VAA *smartvote* – about half of the users would have received a different recommendation. Additional analyses showed that these findings did not stem from small differences between the best and second-best match, but from more substantial differences in the perceived match under different methods. So the degree of match between voter and party as calculated by a VAA depends not only on the positions both take on the policy items in the test, but also on the method used to calculate the match.

The differences at the individual level did not cancel each other out at the aggregate level. On the contrary. The number of 'best matches' presented for each party varied across the methods. For example, whereas 19 per cent of the users received Labour Party as advice in the original application, in a two-dimensional model this figure would drop to 3 per cent. And whereas 3 per cent of the users of *StemWijzer* received Christian Union as voting recommendation, in a two-dimensional model this figure would increase to 28 per cent. In theory it is possible that these findings are (partially) caused by the nature of the statements of *StemWijzer*, which were not selected with a spatial model in mind. However, we believe that the results are inherent to the use of low-dimensional spatial models in VAAs. This is confirmed by findings presented in an unpublished paper that analysed user answers from the 2006 edition of *Kieskompas*, which adopted a two-dimensional spatial framework (Kleinnijenhuis and Krouwel, 2008, Table 2). They found that on the 36 statements in this test 21 per cent of the users agreed most often with one of the two largest Christian parties (CDA and Christian Union). However, in the two-dimensional spatial model adopted by *Kieskompas*, 62 per cent of the users received a recommendation favouring one of both parties. Another 21 per cent of the users agreed most often with the Socialist Party, but based on the two-dimensional model only 2 per cent of the users received the advice to vote for this party. These figures

support our conclusion that the use of a spatial model as such has a strong impact and accounts for differences in advice observed across methods.

To some extent the differences between the outcomes obtained under alternative methods may stem from other design issues than the spatial model. The fact that *StemWijzer* selected its statements in such a way that any pair of parties can be adequately distinguished may have limited the ability to map policy preferences in a low-dimensional spatial model. Furthermore, by encouraging or even forcing parties to take position either in favour or against an issue statement, parties may be clustered more towards the end of policy dimensions than they otherwise would have been (users chose the 'neither' option in 11.2 per cent of cases, while the parties only selected it in 2.4 per cent of cases). It is therefore likely that the scaling problems we encountered will be alleviated, at least to some extent, by registering voters' and parties' positions on statements on a 5-point or 7-point scale. This might also increase the match between low-dimensional and high-dimensional models. However, the main problem appears to be that policy preferences of voters are not strongly structured and simply cannot be captured by any low-dimensional spatial model (Kleinnijenhuis and Krouwel, 2009). This has been a classic issue in political science, since Converse (1964) claimed that the ideological constraint among the mass public is rather limited.

How can these findings be taken into account when VAA makers choose the method for calculating and presenting the advice? First, the findings have an important implication for the way that the advice is presented. In the election voters can usually only select one party or candidate and their decision making is therefore facilitated most by providing the name of a single party, the 'best match'. However, the downside is that this may suggest more precision and objectivity than what can be substantiated. So presenting the full rank ordering of parties with the degree of agreement reflected in a bar chart or a two-dimensional spatial model in which user and parties are represented, as is often done, are more appropriate ways to present the advice than merely providing the name of a single party.

The second and arguably most important implication of the findings concerns the spatial method that VAAs adopt. If the aim is to inform users about the degree to which their own policy preferences match with the positions taken by the competing political parties (or candidates), and if we take the statements included in the test as a given, then the high-dimensional models are superior. If the results of high-dimensional and low-dimensional models would be similar, the latter might be preferred because they additionally provide insight in the main dimensions of political competition. However, the results of VAAs become very different when low-dimensional models are applied. This means that users get the impression that a particular party provides the best match, whereas there is another party with which they agree more often. How would a VAA maker be able to convincingly explain to users that they may agree with a particular party on 75 per cent of the statements, but it still better to vote for another party with which they agree only 50 per cent of the time? Or how would a VAA maker be able to explain to a political party that

30 per cent of the users agree with them most often, but only 5 per cent receives the recommendation to vote for them? These examples capture the essence of the situation. The analysis has shown that these are not just hypothetical examples, but that such situations occur in reality. Therefore, the use of a low-dimensional model to determine the match between voters and parties is problematic.

If VAA designers would still favour the use of a low-dimensional spatial model, they should check whether the model they consider is indeed a valid depiction of the party competition and whether it makes sense as a tool to position voters (Otjes and Louwerse, 2011). This means that they should explore the metric properties of the scales that are constructed with the items included in the VAA, and preferably make these findings public. Such analyses could be incorporated in the design process and help to select suitable statements. If policy preferences of citizens are constrained, and if the individual items from a test form reliable scales of underlying policy dimensions, the use of a spatial model is warranted and valuable. If these conditions are not met, it is better if VAA designers refrain from using low-dimensional spatial models.

To conclude this article, let us emphasise that we have focused on only one element of the design of VAAs, namely the use of a spatial framework and metric to calculate the match between voters and parties on the basis of a given set of items. The quality of the VAA as a whole depends on all elements, including the selection of appropriate topics, adequate phrasing of statements, correct coding of party positions, and much more. We hope that other scholars feel encouraged to explore such issues, in order to evaluate the tools that so many voters use and enable the designers to make wise decisions when deciding about the design of future editions.

## About the Authors

Tom Louwerse is Assistant Professor in political science at the Department of Political Science, Trinity College, Dublin. His research interests include political representation, legitimacy, political parties and parliamentary politics.

Martin Rosema is Assistant Professor of political science at the University of Twente. His main area of expertise is the psychology of voting. He has published on topics such as partisanship, strategic voting, and electoral turnout. Other themes in his current research are the quality of political representation, the role of emotions, and voting advice applications.

## Acknowledgements

# Notes

1 Obviously, some voters may have filled out the VAA multiple times, but this figure does give some idea of the popularity of the *StemWijzer* in the Netherlands.

2 In earlier days its name was *Stichting Burgerschapskunde* and *Instituut voor Publiek en Politiek*. The present name in full is *ProDemos: Huis voor Democratie en Rechtsstaat*.

3 The information about *StemWijzer* in this section is largely based on a chapter by De Graaf (2010). The same method has been applied in Germany, as discussed by Marschall and Schmidt (2010). The user statistics reported, as well as the log files of the 2010 edition, have been provided to us by the developer of *StemWijzer*, Instituut voor Publiek and Politiek in Amsterdam. We are grateful to Jochum de Graaf for arranging this.

4 These figures concern the number of times that the online test was taken. Presumably, there were individuals who completed the test more than once, so the number of voters who took the test may be somewhat lower than these figures suggest. However, data from the *Dutch Parliamentary Election Studies* 2006 and 2010 indicate that in both years about 40 per cent of those who cast their ballot had consulted a VAA.

5 Some earlier versions of *StemWijzer* made use of Euclidean distance, but at some point this was changed into city block distance. *Source*: personal conversation with Jochum de Graaf from Instituut voor Publiek en Politiek, Amsterdam, 25 May 2011. See also Marschall and Schmidt (2010).

6 The precise method used to calculate the agreement scores is not described by Lobo *et al* (2010). We presume that distances between voter and party on each individual item were calculated and then added for all 28 items, leading to a scale ranging between 0 (full agreement on all items) and 112 (maximum disagreement on all items: $28 \times 4$).

7 Not all of the statements are included in the *smartspider* model, while some statements are assigned to more than one *smartspider* dimension. The voter and party positions on each of the dimensions are calculated by using a city block model.

8 This is a very low threshold and only serves to exclude items that do not at all relate to any of the dimensions. If we would apply the more commonly used $R^2$ threshold of 0.8 very few items would load on the dimensions at all.

# References

Aarts, K. and Thomassen, J. (2008) Dutch voters and the changing party space 1989–2006. *Acta Politica* 43(2/3): 203–234.

Cedroni, L. and Garzia, D. (eds.) (2010) *Voting Advice Applications in Europe: The State of the Art*. Napoli, Italy: Scriptaweb.

Converse, Ph. E. (1964) The nature of belief systems in mass publics. In: D. Apter (ed.) *Ideology and Discontent*. New York: Free Press, pp. 206–261.

De Graaf, J. (2010) The irresistible rise of Stemwijzer. In: L. Cedroni and D. Garzia (eds.) *Voting Advice Applications in Europe: The State of the Art*. Napoli, Italy: Scriptaweb, pp. 35–46.

Downs, A. (1957) *An Economic Theory of Democracy*. New York: Harper & Row.

Dumont, P. and Kies, R. (2012) Smartvote.lu: Usage and impact of the first VAA in Luxembourg. *International Journal of Electronic Governance* 5(3/4): 388–410.

Enelow, J.M. and Hinich, M.J. (1984) *The Spatial Theory of Voting: An Introduction*. Cambridge, UK: Cambridge University Press.

Fivaz, J. and Nadig, G. (2010) Impact of voting advice applications (VAAs) on voter turnout and their potential use for civic education. *Policy & Internet* 2(4): 167–200.

Fivaz, J. and Schwarz, D. (2007) Nailing the pudding to the wall – E-democracy as catalyst for transparency and accountability. Paper presented at the International Conference on Direct Democracy in Latin America, Buenos Aires, Argentina.

Garzia, D. and Marschall, S. (2012) Voting advice applications under review: The state of research. *International Journal of Electronic Governance* 5(3/4): 203–222.

Gemenis, K. (2013) Estimating parties' policy positions through voting advice applications: Some methodological considerations. *Acta Politica* 48(3): 268–295.

Kleinnijenhuis, J. *et al* (2007) *Nederland vijfstromenland: De rol van de media en stemwijzers bij de verkiezingen van 2006*. Amsterdam, the Netherlands: Bert Bakker.

Kleinnijenhuis, J. and Krouwel, A. (2008) Simulation of decision rules for party advice websites. Paper presented at the Annual Meeting of the American Political Science Association, Boston.

Kleinnijenhuis, J. and Krouwel, A. (2009) Dimensionality of the European issue space. Paper presented at the NCCR Workshop Vote Advice Applications, Bern.

Ladner, A., Felder, G. and Fivaz, J. (2010) More than toys? A first assessment of voting advice applications in Switzerland. In: L. Cedroni and D. Garzia (eds.) *Voting Advice Applications in Europe: The State of the Art*. Napoli, Italy: Scriptaweb, pp. 91–123.

Ladner, A., Fivaz, J. and Pianzola, J. (2012) Voting advice applications and party choice: evidence from smartvote users in Switzerland. *International Journal of Electronic Governance* 5(3/4): 367–387.

Lobo, M.C., Vink, M. and Lisi, M. (2010) Mapping the political landscape: A vote advice application in Portugal. In: L. Cedroni and D. Garzia (eds.) *Voting Advice Applications in Europe: The State of the Art*. Napoli, Italy: Scriptaweb, pp. 143–171.

Marks, G., Hooghe, L., Nelson, M. and Edwards, E. (2006) Party competition and European integration in the east and west: Different structure, same causality. *Comparative Political Studies* 39(2): 155–175.

Marschall, S. and Schmidt, C.K. (2010) The impact of voting indicators: The case of the German Wahl-O-Mat. In: L. Cedroni and D. Garzia (eds.) *Voting Advice Applications in Europe: The State of the Art*. Napoli, Italy: Scriptaweb, pp. 65–90.

Mendez, F. (2012) Matching voters with political parties and candidates: An empirical test of four algorithms. *International Journal of Electronic Governance* 5(3/4): 264–278.

Otjes, S. and Louwerse, T. (2011) Spatial models in voting advice applications. Paper presented at Politicologenetmaal; 9–10 June, Amsterdam.

Pellikaan, H., Van der Meer, T. and De Lange, S.L. (2003) The road from a depoliticized to a centrifugal democracy. *Acta Politica* 38(1): 23–49.

Ramonaitė, A. (2010) Voting advice applications in Lithuania: Promoting programmatic competition or breeding populism? *Policy & Internet* 2(1): 117–147.

Ruusuvirta, O. and Rosema, M. (2009) Do online vote selectors influence electoral participation and the direction of the vote? Paper presented at the ECPR Conference, Potsdam, Germany.

Trechsel, A.H. and Mair, P. (2011) When parties (also) position themselves: An introduction to the EU profiler. *Journal of Information Technology & Politics* 8(1): 1–20.

Wagner, M. and Ruusuvirta, O. (2012) Matching voters to parties: Voting advice applications and models of party choice. *Acta Politica* 47(4): 400–422.

Walgrave, S., Nuytemans, M. and Pepermans, K. (2009) Voting aid applications and the effects of statement selection. *West European Politics* 32(6): 1161–1180.

Walgrave, S., Van Aelst, P. and Nuytemans, M. (2008) 'Do the vote test': The electoral effects of a popular vote advice application at the 2004 Belgian elections. *Acta Politica* 43(1): 50–70.

Wall, M., Krouwel, A. and 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 Politics*, advance online publication 7 March, doi: 10.1177/1354068811436054.

Wall, M., Sudulich, M.L., Costello, R. and Leon, E. (2009) Picking your party online – An investigation of Ireland's first online voting advice application. *Information Polity* 14(3): 203–218.

Van der Brug, W. and Van Spanje, J. (2009) Immigration, Europe and the 'new' cultural dimension. *European Journal of Political Research* 48(3): 309–334.

Van der Eijk, C. and Niemöller, K. (1983) *Electoral Change in the Netherlands: Empirical Results and Methods of Measurement*. Amsterdam, the Netherlands: CT Press.

# Appendix A

**Table A1:** Statements in *StemWijzer* 2010

| | % agree | % disagree | % extra weight | 2D party | 3D party | 3D voter | 7D |
|---|---|---|---|---|---|---|---|
| The government has to cut spending with billions. In 2015 at latest the national budget deficit should have been resolved. | 41 | 36 | 19 | 1 | 1 | 3 | 2 |
| People who earn very much should pay more taxes than they do now.* | 59 | 33 | 27 | 1 | 1 | 3 | 1 |
| The amount of child benefits should become dependent on someone's earnings.* | 70 | 26 | 24 | 1 | 1 | 3 | 6 |
| Shopkeepers should be able to determine themselves if they open their shop on Sunday.* | 76 | 20 | 10 | 2 | — | — | 7 |
| People who become unemployed receive unemployment benefits. The amount should be increased in the first months, but the total period of benefits should be strongly reduced. | 44 | 39 | 19 | 1 | 3 | 3 | 6 |
| From the age of 65 onwards everybody receives elderly benefits. This should remain like it is.* | 71 | 22 | 33 | 1 | 2 | 3 | 6 |
| It should become easier for employers to fire employees. | 30 | 58 | 20 | 1 | 3 | 3 | 1 |
| All Islamic schools in the Netherlands should be closed. | 34 | 51 | 13 | 2 | 2 | 1 | 4 |
| Extra money is needed for education and therefore the government should abolish student funding. From now on students should borrow money for their study. | 19 | 72 | 26 | 1 | 3 | 2 | 6 |
| The government pays for public television channels 'Nederland 1, 2 and 3'. At least one of the channels should disappear. | 50 | 38 | 8 | 1 | 1 | 1 | 2 |
| The government should strongly reduce subsidies for art and culture. | 55 | 32 | 11 | 1 | 1 | 1 | 2 |
| In the Netherlands no new mosques may be built. | 31 | 53 | 14 | 2 | 1 | 1 | 4 |
| Someone who owns a house receives tax benefits. This arrangement should remain like it is. | 67 | 23 | 39 | 1 | 1 | 1 | 1 |
| Too many people with high income live in social housing. Therefore they should pay more rent for the same house than people with lower income.* | 57 | 36 | 15 | 1 | 2 | 2 | 1 |

**Table A1:** (*Continued*)

| | % agree | % disagree | % extra weight | 2D party | 3D party | 3D voter | 7D |
|---|---|---|---|---|---|---|---|
| The government should spend extra money on new and wider motorways. | 46 | 38 | 13 | 2 | 1 | 1 | 5 |
| Possession of cars should become cheaper and use of cars more expensive. Therefore a mileage tax should be introduced.* | 43 | 45 | 20 | 1 | 1 | 2 | 5 |
| Everybody should be an organ donor, unless written objection has been made.* | 56 | 37 | 18 | 2 | 1 | — | 7 |
| Assistance with suicide should be possible under specific conditions.* | 83 | 11 | 18 | 2 | 2 | — | 7 |
| There should be more competition between care institutions (free market system). | 35 | 51 | 12 | 1 | 3 | 1 | 1 & 6 |
| The queen has several political tasks. This should be changed, so she has only ceremonial tasks.* | 36 | 49 | 8 | 2 | 2 | 2 | — |
| To give voters more influence they should have the possibility to reverse decisions from national politics by referendum.* | 59 | 28 | 12 | 2 | 1 | 3 | — |
| People who are employed by the government should not be allowed to wear head scarves at work. | 47 | 43 | 14 | 2 | 2 | 1 | 4 |
| There should be much more severe punishment for people who use violence. | 88 | 6 | 35 | 2 | 1 | 1 | 3 |
| The government should make the growth and sales of soft drugs legal.* | 48 | 39 | 11 | 2 | 1 | 2 | 3 |
| Organisers of sports events should pay themselves for the use of police force.* | 57 | 30 | 13 | 1 | 2 | 2 | — |
| The Netherlands should spend much less on developmental aid. | 54 | 33 | 19 | 1 | 1 | 1 | 2 |
| The government should strongly cut spending on defense.* | 53 | 28 | 14 | 1 | 1 | 2 | — |
| There should be a prohibition on very large stables for livestock (mega stables).* | 54 | 26 | 10 | 2 | 1 | 2 | 5 |
| To improve the environment, taxes on meat should be strongly increased.* | 22 | 65 | 8 | 1 | 1 | 1 | 5 |
| The Netherlands should be allowed to build a second nuclear plant. | 42 | 43 | 13 | 2 | 1 | 3 | 5 |

*Note*: Statements marked with * are considered left-wing positions in the one-dimensional model, the other statements are considered right-wing. Columns 3–7 indicate to which dimension each statement belongs.

## Appendix B

### Calculation of voting advice according to the different methods

*Agreement method*

$$A_{rs} = \sum_{i=1}^{30} w_{ri} a_{rsi}$$

where $a_{rsi}$ equals 1 when user $r$ and party $s$ provided the same answer to statement $i$, 0 otherwise, and $w$ equals 2 when the user put extra weight on an issue and 1 otherwise, with $i = 1$ to 30 statements.

*Note*: To show that the agreement model is in fact a high-dimensional model, consider that if we rewrite the model in terms of disagreement $d$:

$$d_{rsi} = |k_{si} - k_{ri}| + |l_{si} - l_{ri}| + |m_{si} - m_{ri}|$$

where $k$, l and $m$ equal 1/2 when the party $s$ or user $r$ agree ($k$), disagree ($l$) or are neutral ($m$) on statement $i$. Thus, when a user agrees with a statement: $k_{ri} = 1/2$, $l_{ri} = 0$, and $k_{ri} = 0$. If party and user do not take the same position, they will differ on two of the parameters $k$, l and $m$, yielding an absolute distance of 1 (for example, when the party agrees and the user is neutral: $|(1/2) - 0| + |0 - 0| + |0 - (1/2)| = 1$). If they do take the same position, the absolute distance is 0. Note that we do not need three parameters $k$, $l$ and $m$, because if $k = 0$ and $l = 0$ it follows that $m = 1$. We can re-parameterise the model by stating $P_1 = k - l$ and $P_2 = m$. This yields the following formula:

$$A_{rs} = \sum_{i=1}^{30} w_{ri}(|p_{1si} - p_{1ri}| + |p_{2si} - p_{2ri}|)$$

or, equivalently,

$$A_{rs} = \sum_{i=1}^{30} w_{ri} \sum_{j=1}^{2} |p_{jsi} - p_{jri}|$$

which is similar to a 60-dimensional city block model (see below).

*City block distance method*

$$CB_{rs} = \sum_{i=1}^{30} w_{ri} |p_{ri} - p_{si}|$$

where $p_{ri}$ is the voters' position on issue $i$ and $p_{si}$ the party's position. The weight $w_{ri}$ equals 2 when the user put extra weight on an issue and 1 otherwise, with $i = 1$ to 30 statements.

*Euclidean distance method*

$$E_{rs} = \sqrt{\sum_{i=1}^{30} w_{ri}(p_{ri} - p_{si})^2}$$

where $p_{ri}$ is the voters' position on issue $i$ and $p_{si}$ the party's position. The weight $w_{ri}$ equals 2 when the user put extra weight on an issue and 1 otherwise, with $i = 1$ to 30 statements.