

**SEVENTH FRAMEWORK PROGRAMME
THE PEOPLE PROGRAMME
MARIE CURIE ACTIONS – NETWORKS FOR INITIAL TRAINING (ITN)**

**ELECDEM
TRAINING NETWORK IN ELECTORAL DEMOCRACY
GRANT AGREEMENT NUMBER: 238607**

**Deliverable D18.1 – New Technologies and Political
Preferences: The Electoral Compass**

Final Report

Experienced Researcher (ER)

Dr. Matthew Wall

Host Institution: Kieskompas



Abstract

This report details the research activities undertaken by the author (Dr. Matthew Wall) during his tenure as an 'Experienced Researcher' (ER) at Kieskompas in Amsterdam, the Netherlands, under the auspices of the Marie Curie Initial Training Network in Electoral Democracy (ELECDEM). The research undertaken in this period focused on data produced by online 'Voter Engagement Applications' (VEAs), also known as 'Vote Advice Applications' (VAAs). The research undertaken was primarily quantitative in nature, and deployed a mixture of survey based and experimental analysis. The research focused on three key questions: examining the impact of VEA sites on the voting behaviour of their users; investigating the reliability and design features of VEA sites to examine their contribution to electoral democracy; and experimenting with methodological innovations made possible by the novel data collection techniques facilitated by databases of VEA users. The report indicates that VEA sites do exert and influence on their users; that design issues are extremely important for the functioning and legitimacy of such sites; and that user databases can be used to contribute to the methodological debate on optimising online survey implementation.

Table of Contents

Page 4. Tables

Page 5. Figures

Page 6. Executive Summary

Page 9. Full Report – 1. Objectives and Aims

Page 16. Full Report – 2 Situating the Research Questions in the Literature

Page 38. Full Report – 3. Measurement, Data and Methods

Page 50. Full Report – 4. Results and Analysis

Page 89. Full Report – 5. Summary and Conclusions

Page 92. References

Tables.

Page 54. Table 1. Responses of survey group to the question: 'Which of the following best describes the influence that visiting kieskompas.nl had on your vote choice?'

Page 56. Table 2. Was VAA advice 'followed'? Congruent versus non-congruent advice (column percentages in parentheses).

Page 58. Table 3. Distributions of 'PTV' for recommended versus not recommended voter-party dyads.

Page 60. Table 4. Proportions of 'voted' user-party dyads for 'recommended' and 'not recommended' dyads for each value of PTV (*N* for each proportion in parentheses, statistically significant differences in boldface).

Page 63. Table 5. Types of MAUD rules explored in this report.

Page 64. Table 6. Aggregate Partisan Recommendations generated by Alternative MAUD Rules (% of first choice recommendations per party).

Page 66. Table 7. Pedersen index of volatility for pairwise comparisons of 7 MAUD rules

Page 68. Table 8. Proportion difference in first-preference vote advice, by party, compared to reported (2006) voting behaviour of site respondents.

Figures.

Page 51. Figure 1. Advice screen of Kieskompas.nl 2010 Dutch legislative elections (closest-party advice circled).

Page 80. Figure 2 – Output displaying text justifying a single issue coding (note hyperlink to source document at the bottom of this text). Example taken from Mexican 2012 Vote Compass: <http://www.brujulapresidencial.mx>

Page 84. Figure 3 – An example of single issue output on the question ‘Richer people should pay more taxes than poorer people’ (translation); with voter ‘no opinion’. Example taken from Mexican 2012 Vote Compass: <http://www.brujulapresidencial.mx>

Page 88. Figure 4. Response rates for each group with 95% bootstrapped confidence intervals.

Page 86. Figure 5. Item-missing count with 95% confidence intervals for all groups.

Executive Summary

‘Voting Advice Application’ (VAA) or ‘Voter Engagement Application’ (VEA) websites are among the most significant online political innovations that have emerged to date in established democracies. Typically, such sites are established during electoral campaigns; they elicit information from users regarding their policy preferences, compare these to the policy positions of political parties or candidates, and issue advice to the user as to which party’s policies best match the user’s preferences. This advice may be presented as a simple discrete recommendation for the user to vote for a specific party, or as an explicit rank ordering of parties or a (two dimensional) visualization of parties relative to the user according to ‘closeness’ to the issue positions of the voter. Guidance may also be given by a graphical presentation of the position of the voter in an n -dimensional (usually two-dimensional) political landscape relative to the issue positions of the various parties.

In this report, I explain the research that I conducted during a one year postdoctoral research fellowship as an ‘Experienced Researcher’ (ER) under the auspices of the ELECDDEM Marie Curie Initial Training Network (ITN) at Kieskompas, a company based in the Free University Amsterdam, founded by Dr. André Krouwel (my academic supervisor during the postdoc). Kieskompas is a Dutch company that began by making a VAA for the 2006 Dutch legislative elections, but which subsequently expanded its operations – offering VEA websites in an ever-growing range of national elections. I begin the report by elaborating the aims and objectives of the research that I carried out – explaining exactly what Kieskompas does and how its operations generate data that is of interest to political science scholars. I explain how the key research questions that I arrived at in the first part of my research period focus on issues that are consonant with the themes and concerns of the wider ELECDDEM project. I outline the three key research questions that drove my research during my fellowship at Kieskompas, and, for the remainder of the report, each section goes through how I addressed these questions separately.

- Key Question 1: Are users of VAA/VEA websites influenced by the ‘advice’ that they receive online? Specifically, are they more likely to vote for a party that they have been ‘advised’ to vote for online?

This was the primary question that I investigated in during my fellowship. In the report I explain how I developed an innovative research strategy that involved collecting data from VEA site users at multiple time points to get a better feel than any existing research on the causality of a VEA ‘effect’. In short, I found convincing evidence that there is such an effect – specifically I find that users are significantly more likely to vote for parties that are recommended to them during their visit to a VEA site – controlling for their prior self-estimated likelihood to vote for that party. However, the research indicates that this effect is not indiscriminate – instead it is only observable when the VEA advises its user to vote for a party that they were already seriously considering as a vote choice.

- (Key question 2) How does the design and execution of VEA sites influence their contribution to both the accountability and the legitimacy of elections? How can VEA site design be improved to maximise this contribution?

I address this question from two perspectives in this report. Firstly, I report on research that focused specifically on the ‘reliability’ of VEA outputs. In this analysis, I discuss the various competing theories that exist in the political science literature that allow for a systematic, overall conception of the extent to which voter and party positions can be said to be ‘similar’ or ‘close’. I demonstrate that each of these conceptions gives rise to an alternative potential ‘Multi-Attribute-Decision-Utility’ (MAUD) rule for matching voters and parties. As all VEAs require a MAUD rule – it is important to investigate whether advice varies dramatically according to the rule that is selected. I find that MAUD rule selection does substantially effect VEA outputs, on the basis of a simulation exercise using real party and voter data from a Kieskompas VEA. I then move to a broader description of the process of designing a VEA – demonstrating that this is a multi-stage process with important choices to be made at each stage. I conclude that VEA designers must seek to maximise transparency at every stage of site implementation – and argue that the uncertainty of the ‘advice’ that

they give (given variance in output according to MAUD rule) means that designers should be careful not to overstate the certainty of their findings. In this way, the contribution of VEA sites to the legitimacy and accountability of the electoral process can be maximised.

- (key question 3) What can data gleaned from Kieskompas email surveys tell us about online survey response rates?

Under this heading, I report the results of experimental work that used a mailing list of Kieskompas users in order to analyse the effects of online survey design on the response rates that such surveys obtain. Specifically, I and my collaborators were interested in investigating the effects of the formulation of the email that accompanies the request for participation. We sought to develop letter wordings that captured differences in tone, complexity, and they type of appeal (egoistic versus altruistic) contained in the request to participate. We found that simple letters, which appeal to altruistic motivations of respondents produced higher response rates. In the conclusion to this report, I go further in examining other potential uses for the highly promising data that are produced by VEA sites.

Generally, the research presented here seeks to better understand the role of VEAs, which, I argue are a feature of the contemporary globalised communications environment, to the quality and functioning of electoral democracy. I find that they are influential, and have a capacity to contribute to the accountability and legitimacy of the electoral process, but that to maximise this contribution it is vital to be conscious of the various choices and pitfalls involved in VEA design. Finally, I argue that VEAs are also important methodological tools for political science scholars involved in electoral studies, and that as scholars find new uses for VEA data, and techniques for applying these data, VEAs will come to be a key tool for the investigation of electoral democracies around the world.

Full Report

1. Objectives/Aims

The objectives and aims of my one year Experienced Researcher (ER) postdoctoral research project were initially laid out in the formal workpackage 18 description, which reads as follows:

‘This workpackage, the responsibility of KIESKOMPAS, utilises data on user issue preferences and party positions. These data will be analysed using advanced statistical techniques such as multilevel modelling. Tasks: 1. Develop literature review, research question and research design; 2. Analyse available data employing advanced quantitative techniques and experimental techniques 3. Issue final report; 4. Produce research paper for submission; 5. Undertake training in key methodologies and complementary skills’.

Before describing how these objectives and aims were refined into a concrete research design and suite of research questions, it is important to provide some explanatory discussion as to what Kieskompas does, and how it generates public opinion and party position data.

What does Kieskompas do?

Kieskompas.nl is a company that produces what are commonly referred to as either Vote Advice Application (VAA) or Voter Engagement Application (VEA) websites (these terms will be used interchangeably in this report). The idea behind a VAA site is that a user visits the site, responds to several questions on political issues, and then receives personalised advice as to how their issue preferences compare to the policy stances taken by parties/candidates

competing in an upcoming election. Such sites are among the most popular directly political applications of internet technology developed, and frequently attract millions of users during election campaigns in a growing number of established democracies (Walgrave et al., 2008). VAA sites are also the most widely consumed political science outputs in the world today, as they either implicitly or explicitly use theories, methods, and data derived from the work of political scientists.

As well as producing websites for local and national elections in the Netherlands, Kieskompas works with academic teams in a wide range of countries to produce VEA sites for their national-level elections. Among the projects that I worked on during my time at Kieskompas were sites in: Sweden; Netherlands; Canada; Turkey; Egypt; Tunisia; and Morocco.

How does Kieskompas generate data?

Kieskompas generates three types of data:

1) Party positioning data:

All parties are positioned on the key issues in each election (See below for a description of how these issues are identified for each site). Parties are positioned on the basis of a hierarchy of texts indicating party policy (usually the election manifesto is the top-ranked source for these analyses).

2) Public opinion data on opinions of political issues:

Members of the public who access VEA websites input their responses to questions on the key issues (these responses are the basis upon which individual voters are compared to the parties competing in the election). Users' responses are registered in a site 'log file' and this log file is thus the source of the public opinion analysis performed by Kieskompas.

3) Data on user demographics and voting behaviour:

Kieskompas also collects 'extra' data from users via 'opt in' surveys that are suggested to users while they are visiting the site. Furthermore, a sub-sample of users (who provide email details and indicate a willingness to be contacted) participate in 'follow-up' surveys which are conducted via email.

In seeking to deliver my workpackage, I was therefore confronted with data that was highly useful for investigating politics and public opinion but also highly original and unusual – meaning that considerable thought went into producing publishable analysis. In the next section, I explain how I achieved step 1 of the workpackage – I was already familiar with the literature in this area, having participated in the development of an Irish VEA site in 2007, and having published a paper based on an analysis of the data derived from that project. However, with greater time resources to dedicate to the study of this topic, and the availability of data from Kieskompas and the expertise of Dr. André Krouwel (my academic supervisor during my time at Kieskompas) I was able to consider broader and more ambitious research questions than was the case in my previous investigations.

1 (a) Conceptual Linkages – Evaluating Electoral Democracy

The concrete aims and objectives that I arrived at in the first stage of my work at Kieskompas were designed to be consonant with the overarching themes of the ELECDEM project.

I sought to focus upon the themes of globalisation and communication and how they are influencing the contemporary realisation of electoral democracy. VEA sites are among the clearest instances of the Internet, which is both a globalised phenomenon and which is arguably the most significant recent development in human communication, intervening in

the conduct of political campaigns and elections. Specifically, I was interested in the following research question, which formed a major theme of my work at Kieskompas:

(key question 1): Are users of VAA/VEA websites influenced by the ‘advice’ that they receive online? Specifically, are they more likely to vote for a party that they have been ‘advised’ to vote for online?

A core goal of the ELECDDEM project is to evaluate how globalisation, communications and institutions contribute to the quality or functioning of electoral democracy. While this workpackage was not well-positioned to examine the role of institutions, I took the VEA sites as a globalised communications phenomenon and investigated how their execution and function contribute to electoral transparency and how such sites should be implemented to maximise their usefulness and trustworthiness for voters.

(Key question 2) How does the design and execution of VEA sites influence their contribution to both the accountability and the legitimacy of elections? How can VEA site design be improved to maximise this contribution?

Finally, as outlined above, VEA sites provide highly unusual data on public opinion and party positions. A key challenge of this project was to develop both practical survey design and implementation ideas, as well as statistical techniques for the analysis of these data. In order to address key question 1 – a highly innovative approach to data collection, collation and analysis was adopted, as shall be explained. However, the data also provide novel opportunities to explore methodological issues that are key to social scientific investigation of the function of electoral democracies. Thus, a final research question asks:

(key question 3) What can data gleaned from Kieskompas email surveys tell us about online survey response rates?

Thus, during my year as an ER at Kieskompas, my research was directed at responding to these three key questions, which tie directly into the core themes investigated in the ELECDEM project. As such, for the remainder of the report, each section will be subdivided into how each of these key questions was addressed. I wish to note here that while, for presentational purposes, I refer to 'my' research work in this report, the research was conducted in close collaboration with Dr. Krouwel throughout – I also worked collaboratively with two other ELECDEM scholars – Zoltan Fazekas and Thomas Vitiello. Furthermore, Dr. Krouwel and I collaborated with Prof. Jan Kleinnijhuis, a Professor of Communication Studies at the Free University Amsterdam (where Kieskompas is based).

1 (b) Key Methodologies and Data

Key question 1 was addressed using, in the first instance, bespoke survey questions developed by myself and Dr. André Krouwel (my academic supervisor at Kieskompas). Later in the project, we were joined by Thomas Vitiello, on an ELECDEM secondment to Kieskompas, and Thomas assisted with the development of our theoretical understanding of how a campaign event can 'influence' a person's eventual vote choice. In our analysis, we relied on original data generated by the 2010 kieskompas Dutch legislative election site, which were supplemented by the bespoke survey questions in a post-election follow up on a sub-section of users (those who had provided an email address and indicated that they were willing to participate in a follow-up). As I explain below, linking these two sources of data on the same voters proved vital to the success of the research design for addressing key question 1. In designing these questions I relied on the training that I received during my participation in the ELECDEM 'Electoral Survey Design' (ES) workshop in Budapest in April, 2010. Having designed the relevant survey questions I compiled the resulting data and finally tested my research hypotheses using advanced quantitative methods – the specific

methodologies deployed are elaborated in detail in the main body of this report. These methods were developed especially through my participation in ELECDDEM's in June/July 2010 Brussels workshop in Experimental Methods (EX).

Key question 2 was addressed in two discrete research papers. The first paper sought to isolate the effects of the algorithms that VEA sites use – in order to determine the extent to which the eventual results of VEA sites are influenced by choice of algorithm – it relies on advanced quantitative analysis – specifically simulations of VEA advice using real public opinion and party positioning data (again from the Netherlands) which would have been produced under alternative algorithm specifications. Data preparation for this project began with inputting data detailing the positions on 30 key election issues of the parties competing in the 2010 Dutch elections (as measured by Kieskompas – see 'How does Kieskompas generate data, above) into an identical format with the input data of site users, who entered their opinions on these same issues when visiting the site. Then, specific algorithms, reflecting competing theoretical descriptions of party policy completion and party-voter 'closeness' were derived. Simulations using these algorithms generated data on how the Kieskompas results would have looked under each algorithm, and the analysis then compares aggregate volatility across different types of algorithms as well as the types of parties that specific algorithms appear to favour.

The second paper addressing key question 2 built on my experiences and on the experience of Dr. Krouwel, including work conducted in Toronto on the Canadian project and work conducted in Istanbul on the Turkey project, as well as my learnings from participating in several international conferences – including two sessions of the Dutch Political Science Association's Annual Conference and the European Consortium for Political Science's General Conference - with other VEA designers. Again, we worked with Thomas Vitiello, who gained considerable experience of the ins and outs of VEA design while working on a large scale French VEA project. We provided a qualitative review of the 'kieskompas method' for implementing VEA, reflecting on how this method enhances the that a VEA site can make to the legitimacy of the electoral process. Many of the insights and methodologies

discussed in this paper were developed and/or refined at the ELECDEM workshop on Content Analysis (CA) held in Amsterdam in March, 2011. The paper also employs descriptive data analysis from an EU-wide VEA project.

The third key question was addressed using data from a follow-up survey on a sub-section (those who had filled in their email addresses and given permission to receive future surveys) from the Kieskompas Netherlands 2010 website. In this case, data collection was highly complex, as the follow up to the survey was divided into several experimental groups, with each group receiving a specific treatment. The treatment that we developed was the type of email that respondents receive – we were interested in measuring whether variation in tone, complexity, and the type of appeal made to respondents influenced response rates across experimental groups. In designing these treatment letters, and particularly in examining whether the resultant letters represented valid operationalisations of these concepts of interest, we relied heavily on quantitative textual analysis techniques that we were introduced to in the ELECDEM Workshop in Advanced Techniques for Communication Research: Content Analysis (CA) held in Amsterdam in March, 2011. Again, this research relied on quantitative analysis developed by myself, Dr. Krouwel and ELECDEM Early Stage Researcher (ESR) Zoltan Fazekas during ELECDEM's ES and EX workshops. Specifically, we designed survey questions to test the 'quality' of survey response, examining response consistency when identical items were 'reversed' in the survey. In order to examine the impact of our treatments on response likelihood and quality, we deployed an array of group-comparison and multivariate methodologies in our analysis, these are described in detail below.

2. Situating the Research Questions in the Literature

- Key Question 1: Are users of VAA/VEA websites influenced by the ‘advice’ that they receive online? Specifically, are they more likely to vote for a party that they have been ‘advised’ to vote for online?

VEA sites are a comparatively novel political phenomenon, and consequently political scientists have only recently started to investigate how they function, what sort of advice they produce, and who visits them. Farrell and Schmitt-Beck (2008) argue that VAA sites are instances of a broader trend in the postmodern campaigning environment: the growing presence of non-party actors, who participate in and often shape electoral campaigns, without themselves running for office. VAA ‘campaigns’, however, differ from those of many other non-party actors because they produce voting advice that varies from individual to individual, rather than consistently supporting any one party, and they claim that their recommendations are specifically tailored to each individual’s policy preferences. As such, VAA sites represent a uniquely personalised and directed source of political *information* that voters access during campaigns.

Importantly, in many large countries, including the Netherlands, Germany, France, and Canada – VEA audiences frequently run into the millions. As with internet technologies generally, there are indications that the gap between VEA users and the rest of the population is narrowing over time. This normalisation appears set to continue, as Internet net use becomes more and more diffused in the population, and as VAA use by voters becomes a ‘normal’ part of political campaigns (Fivaz and Schwartz, 2007). The obvious question that therefore arises is whether such sites are political influential.

Walgrave et al. (2008) argue that the potential for VAAs to influence their users lies in their informative effect. They argue that a major function of VAAs is to substantially reduce the cognitive cost needed for a voter to engage in informed issue voting, which was previously

prohibitively costly due to the time and effort required to gather and consider information parties' policy positions, and to compare these positions to one's own policy preferences. This cognitive price tag is particularly exorbitant in highly fragmented, multidimensional party systems such as the Netherlands, so consequently the cognitive savings offered by VAA sites are all the greater. Walgrave et al. conclude that 'the fact that VAAs seem to be popular, especially in countries with a large and fragmented, and thus complicated, party system, indicates that information is key' (p.43).

Public opinion research going back to the earliest uses of scientific public surveys in political science (*e.g.*, Campbell et al., 1960) has consistently grappled with findings that indicate low levels of voter information among populations in established democracies. Lau and Redlawsk (2001: 951), for instance, argue that 'the widespread ignorance of the general public about all but the most highly salient political events and actors is one of the best documented facts in all of the social sciences'. Given the complexity of day-to-day politics at the nation-state level, and the apparent lack of political information possessed by voters, many analysts have asked whether the political opinions of individuals can be considered as reasonable and rational, reviving a debate that dates back to the very origins of democratic theory: how can the people decide on a political direction when they lack the expertise necessary to understand the issues at hand? Do low levels of public knowledge on political issues mean that the public should be proscribed in their capacity to shape public policy by a 'responsible' political elite, because highly participatory democracy 'marginalizes the wise' (Held, 1996: 31)?

A focus on the cognitive processes that operate when individuals make decisions with little information has led several analysts to conclude that voters employ 'heuristics' or 'cognitive short cuts' to arrive at decisions that are rather similar to those that they would have made had they investigated the issue in depth, just as traffic signals allow motorists to drive safely through an intersection without complete knowledge of incoming traffic patterns on all roads leading into that intersection (Lupia, 1994; Lupia and McCubbins, 1998). Such short cuts are necessary if voters wish to make preference-maximizing political choices without the excessive cost in time and energy that complete knowledge demands

and allow voters to enjoy meaningful political preferences. This optimistic interpretation has been challenged by those who believe that candidates, parties, and media actively distort the heuristic signals received by voters, persuading them to vote differently than they may have done with complete information (*inter alia* Bartels, 1996; Lau and Redlawsk, 2001; Vavreck, 2001).

There are numerous potential shortcuts available to voters; Lau and Redlawsk (2001) provide a comprehensive overview of the types of heuristics that may be used by a voter during a campaign. Evidently, affective partisanship or party identification is the most straightforward simplifier of political life, however as such partisanship is in decline (Dalton and Wattenburg, 2000), voters are increasingly forced to look elsewhere for vote cues. Fiorina (1981) proposed that one's personal circumstances and experiences provide all voters with a powerful heuristic for evaluating incumbents. However, in multiparty systems with low identifiability of executive responsibility like the Netherlands, such heuristics may be less easily available. In majoritarian systems, public opinion polls can provide a useful cue as to the electoral viability of individual candidates or parties (Lau and Redlawsk, 2001). This information is less relevant in political systems like the Netherlands which employ highly permissive proportional representation electoral systems, meaning that all but the smallest parties are electorally viable. In this way, political institutions, a key theme of the ELECDEM project, may influence both the popularity and influence of VEA projects, however I was unable to systematically explore this contention in my research at Kieskompas.

Personality-based evaluations of individual candidates or party leaders also offer an important heuristic (Marcus, Neuman, and MacKuen, 2000; Popkin, 1991). Anthony Downs (1957) envisioned ideology as a cost-effective cognitive shortcut which enables voters to understand the broad inclinations of parties without having to learn their detailed policies. Finally, political endorsements by interest groups, or by one's family or peers, also have a powerful heuristic role – as they provide a specific voting recommendation, although individuals face difficult choices when endorsements conflict (Mcdermott, 2006).

Within the nascent field of VEA studies, Ladner et al. (2010) have identified a 'second wave' of VAA studies, which have investigated VAAs by asking at the same fundamental questions of them that political scientists have asked of previous campaign innovations: 'What are their effects on voting behaviour?', and 'How do they relate to the realisation of democracy writ large?'. I focus here on studies that have investigated the influence of VAA use on vote *choice*, that is, the specific party that individuals vote for, rather than their influence on turnout or broad political engagement. For promising research on these topics, which indicates that VAA site use can increase both turnout and political engagement among users, see: Marschall and Schmidt (2010); Ruusuvirta and Rosema (2009).

There are multiple techniques that have been deployed to date in the investigation of VEA effects. The techniques that was developed in my research differs from all of these, however, a review of previous efforts helps to situate my approach in the literature. Several studies investigating the effects of VAAs on their users' vote choices have employed post-election surveys of users (sometimes as part of larger surveys which also include non-users), where respondents provided their own subjective evaluations of whether their choice was influenced by their visit to a VAA site (Carlson and Strandburg, 2005; Aart and van der Kolk, 2007; Ladner et al., 2010; Marschall and Schmidt, 2010; Walgrave et al., 2008). These surveys have varied dramatically in their estimates of the importance attributed by users to VAA sites. A lack of standardisation in the field to date means that the specific questions used to elicit estimates of site influence vary across studies, which may help to explain some of the disparity of findings. Estimates of percentages of users who feel that their eventual decision was influenced by their visit to a VAA vary from a low of 6% (Marschall, 2005) to a high of 67% (Lander et al., 2010).

However, subjective evaluations of the extent to which an event or recommendation was influential after the fact, while informative, are regrettably not totally reliable sources of information as to the actual influence that the event may have exercised. The agenda-setting, priming and framing literatures in communication and media studies, for instance, have uncovered the existence of politically influential behaviours that rarely register in the consciousness of voters (Scheufele and Twexbury, 2007). More generally, post-election surveys provide limited analytical leverage over the impact of any single campaign event on

voting patterns, which is why specific designs, such as survey panels and rolling cross-sections (Bartels, 2006; Brady and Johnston, 2006) have been employed by researchers interested in campaign dynamics.

Several studies have therefore used panel surveys to compare the behaviour of users to non-users of VAAs in terms of their proclivity to 'switch' their first preference vote choice during the campaign and/or to 'switch' party support compared to their previous vote (Ladner et al., 2010; Ruusuvirta and Rosema; 2009; Walgrave et al., 2008). These studies have also generated mixed findings. Walgrave et al. (2008) find little support for the contention that VAA users are any more likely to switch vote intentions than non-users during the campaign, whereas Ladner et al. (2010) and Ruusuvirta and Rosema (2009) both conclude that VAA users are significantly more likely to switch than non-users. Ruusuvirta and Rosema's analysis is particularly insightful, they argue that the content of the advice (confirming or disconfirming existing preferences) has to be considered in investigating VAA effects, and they find that those who reported receiving preference-disconfirming advice when they used a VAA are nearly 3 times more likely to switch vote intention during the campaign than those who received preference-confirming advice.

In spite of these advances, research designs that seek to infer an estimate of the effects of VAA sites on voting behaviour using such data face major difficulties. The first problem is one of causal endogeneity. VAA sites can attract high numbers of unaligned or wavering voters (Ladner et al., 2010). Analyses of whether users of VAA sites exhibit higher in-campaign or between-election volatility than non-users may therefore tell us more about the type of audiences that VAAs attract than about the effects that they may be said to exert. Similarly, if a user receives disconfirming advice, this may be because changes in either the user's opinions or the party's position mean that there are objective reasons why that user should decide to switch parties during the campaign. Again, it is difficult to distinguish instances of a VAA site exerting an influence from those where the site identifies a change in preferences that has already taken place (Ruusuvirta and Rosema, 2009).

Another problem for such analyses is that a visit to a VAA site is not a mass-mediated ‘campaign event’ in the typical sense of the phrase, because the informational output of a VAA site (*i.e.*, the ‘vote advice’) is not identical for each user. As such, even when researchers incorporate panel designs into their studies of VAA users, the dichotomous variable: ‘used/didn’t use a VAA’ is a rather clumsy proxy for each user’s experience. Unfortunately, this is not a problem that can be resolved by incorporating a survey item asking users to recall the specifics of the advice that they received.

My approach in addressing key question 1 (see subsequent sections for details) allowed me to compare users’ recall of the vote advice that they received with objective data from their log files. I find that user recall of advice received is rather unreliable: 53.8% of users recalled receiving advice that was different from the advice recorded in their log files. It is possible, of course that the users visited the site before or after the visit where they left their email address (*i.e.* the visit for which their log file was recorded). This may mean that user ‘error’ figures are slightly inflated. Nonetheless, absent any other evidence, the proclivity of users to ‘recall’ their advice being the party for whom they voted when their visit records indicate that this was not the case, poses a significant methodological problem for user recall-based VAA studies.

Furthermore, it appears that user recall is consistently biased: respondents who did not correctly recall the advice recorded in their log files are far more likely to ‘recall’ the party that they voted for than any other party. Of the over 2,000 incorrect recalls, 749 individuals (34.4% of users who incorrectly recalled the advice received) stated that they had been advised to vote for the party which they had voted for in the election. This observed bias means that using respondent recall to measure VAA advice risks inflating estimates of the extent to which VAA users ‘followed’ the advice that they received. Consequently, a recall-based research design may lead to significant overstatements of the causal importance attributable to VAA site advice. This finding re-enforces Vavreck’s (2007) conclusion that voter recall data can substantially over-estimate the effects of campaign events.

Researchers thus face significant difficulties in estimating the electoral effects of VAA use – given the state of the art at the beginning of my research period with Kieskompas, I therefore felt that this was an area where significant improvements could be made to our understanding of the effects of VEA sites with an innovative approach to data collection and analysis. I argue that the approach outlined in my approach to addressing key question 1 overcomes several of the problems encountered in the extant literature. The combination of log files with pre and post-advice survey items incorporated in my research design provides greater insight than previous studies into the casual process that connects voting advice received online to actual voting behaviour. However, several challenges remain for future studies: I cannot reliably extrapolate an estimate of the importance of VAA sites for the Dutch voting population on the basis of these data, nor can I comment on the extent to which VEA impact varies across countries or elections – a problem that I return to in my conclusions.

- (Key question 2) How does the design and execution of VEA sites influence their contribution to both the accountability and the legitimacy of elections? How can VEA site design be improved to maximise this contribution?

In a context of weakened partisan socialisation (Dalton and Wattenburg, 2002), political scientists have noted the presence in electorates of a growing number of ‘de-aligned’ and ‘non-aligned’ voters who change party preferences between elections (Dalton and Gray 2003; Franklin 1992; 2003). Modern elections take place in a seemingly ever more fragmented and complex media environment, where fact and opinion, policy and spin are difficult to distinguish. In this context the popularity and impact of VEA sites is not entirely surprising. Such sites purport to simplify the intricate and often contradictory statements and signals given to the public by political parties/candidates into an easily understandable representation of the political world, and to allow users to examine where their own preferences place them relative to the parties competing in that space. Their functionality is analogous to that of market aggregator sites in the insurance and travel industries. They bring together the disparate offerings of several groups and provide a ranking and/or recommendation as to which offering best matches the user’s individual preferences. Such

sites also have a certain entertainment value; and some of their popularity is undoubtedly due to the suspense that they can produce as users wait to see which party they most closely resemble.

The value of such sites and the reliability of the advice that they produce to users have been subject to considerable debate in countries where they have become popular. Proponents of VAA sites argue that they have the potential to improve the quality of democratic participation in modern democracies by making it easier for voters to become informed about the policy positions of political parties (Walgrave *et al.*, 2008). A core element in many theoretical and empirical treatments of voting behaviour in modern representative democracies is that voters base their decisions at least partially on an assessment of the likely policy actions that parties will undertake in the future. Indeed, the 'mandate' model of political representation sees this type of consideration on the part of voters as a key source of legitimisation for the exercise of political power in a representative democracy.

Given the costs of analysing parties policy outputs, as well as the noted tendency for parties to 'talk past' each other, seeking to emphasize those issues that they 'own' rather than staking a clear stance on all issues, voters have strong incentives to seek informational shortcuts that allow them to estimate which parties are most likely to act in accordance with their preferences. While, under the classical Downsian (1957) conception, voters were able to use positions on the 'left-right' scale to perform this shortcut, VAA sites potentially allow voters to compare their positions to the main parties on a wide range of issues to their own preferences. According to this conceptualisation, VAA sites represent a public good; a free source of positional information on the main parties/candidates, condensed into an easy-to-understand and, indeed, entertaining online format.

VAA sites also have trenchant critics, who have argued that they unrealistically simplify complex political issues (Hooge and Teepe, 2007). For such critics, VAA sites are, at best, incapable of providing reliable advice and, at worst, politically pernicious devices, which

may give biased advice; systematically favoring certain parties over others, and perhaps even encouraging the unprincipled adoption of populist issue stances by parties and candidates seeking online endorsements (Ramonaitė, 2010). However, the validity and reliability of VAA sites' voting advice have received surprisingly little sustained attention from political scientists. Walgrave *et al.* (2009) point out:

'in sharp contrast to their amazing popularity, to the pertinent questions about VAA outputs, and to the fierce political debate in some countries, the *scientific* debate about VAAs has hardly commenced' (p. 1162).

My research under key question 2 sought to narrow this lacuna in the extant literature, building on the approach adopted by Walgrave *et al.* (2008) and by myself and colleagues in a paper based on data from an Irish VEA project (Wall *et al.*, 2009) by simulating the advice that would have been generated for users of a VAA site using alternative methodologies. While Walgrave *et al.* (2008) focused their attention on the impact of statement selection; one of my principal focuses in addressing key question 2 was on the significance of the *decision rules* used by VAA designers to compare users' policy preferences to parties' issue stances, which represents a significant step forward in this literature.

Regardless of the manner in which party and voter policy preferences are measured, and regardless of the specific issue statements that are used to capture the policy conflicts that characterise the political debate in a given election, all VAA sites must employ a specific Multi-Attribute-Utility-Decision (MAUD) rule in order to give voters an overall indication of how 'close' or 'similar' they are to each of the competing parties in terms of policy preferences. A MAUD rule is simply an algorithm that is used to compare options over several dimensions and to arrive at a decision. In the case of VAA sites, a MAUD rule must compare the opinions of voters with the policy platform of each party incorporated in the website, and then arrive at a decision as to the relative ranking of each party, going from 'closest' or 'most in agreement' to 'furthest' or 'least in agreement'.

MAUD rules embody assumptions with regard to the comparison of user and party stances in order to arrive at a set of scores that measures of the relative utility that each user should associate with each party's policy stances. Any specific MAUD rule for a VAA site can only be built by making a number of decisions about the nature of the issue space within which party issue stances and voter preferences are located. The problem faced by VAA site designers is that when deciding among possible assumptions about the nature of the political issue space and the calculation of closeness/utility in order to choose a MAUD rule, they find themselves confronted with a near-infinite number of possible approaches. Furthermore, there is very little agreement among political scientists concerning some of these fundamental issues. In this review of the relevant literature, I seek to identify some major sources of potential differentiation in how one would go about comparing voters' policy preferences to party policies, and examine their implications for designing a MAUD rule. All of these topics have received more detailed individual attention elsewhere in the literature; here I simply recap the broad parameters of each debate and outline their implications for building a MAUD rule to issue policy-based voter advice.

Proximity Theory versus Directional Theory

Socio-structural explanations for voting behaviour have declined across Europe (Franklin et al., 1992; Dalton and Wattenberg, 1993; Dalton, 1996) and especially in the Netherlands (andeweg and Irwin, 2002: 89-92) leading those seeking to explain electoral behavior to look to issue voting as the key explanatory factor in explaining vote choice in modern electorates (van der Brug, 2004). The conflict between 'proximity' theory and 'directional' theory is foremost among debates in the political science community regarding the manner in which relative voter and party positions translate into electoral utility. These theoretical accounts of vote choice are built on radically different conceptualisations of the very nature of political contestation. The proximity model of electoral choice was introduced in Downs' (1957) seminal work, which transposed Hotelling's (1929) and Coombs'(1950) insights on patterns of spatial competition in business and individual choice behaviour respectively to the study of party policy competition and vote choice. The underlying analogy that drives

this model of politics presented by Downs is that parties' political stances and voters' preferences can be understood in spatial terms as political positions in a hypothetical issue space. In terms of the relationship between voters' policy preferences and parties' policy positions, proximity theory assumes that, all else being equal, voters will prefer the party/candidate whose policy position is closest to their own in that issue space.

Proximity theory views the issue space within which the policy preferences of voters and parties are located as being made up of continuous, graded dimensions; with each dimension containing an infinite number of possible positions between two extreme points. In the simple, uni-dimensional case, the party that will be most-preferred by each voter is the party that has the smallest absolute distance from that voter's ideal policy position on that policy dimension. Furthermore, all other parties can be ranked in terms of their desirability according to their absolute distance from the voter, with distance being inversely related to desirability. This approach is of fundamental to the development of several areas of political science. Even two of its most trenchant critics acknowledged that 'no other formal paradigm has such wide use or such great impact on how people think about politics' (Rabinowitz and MacDonald, 1989: 93). Attempts by political scientists to construct theoretical models of the behaviour of parties when forming coalitions, the actions of parliamentarians, and the behaviour of institutional actors have all relied heavily on the proximity paradigm. Furthermore, proximity theory has been and remains central to political scientists' understanding of policy competition in elections, and of voters' policy-based evaluations of political parties.

Directional theory, first introduced by Rabinowitz and MacDonald (1989) assume that policy alternatives are dichotomous – essentially, voters are either 'for' or 'against' all policy stances, with variations only in the fervour of their support or opposition. Two of the major proponents of directional theory summarize the theoretical assumptions underlying their approach as follows:

‘the directional theory assumes that most people have a diffuse preference for a certain *direction* of policy-making and that people vary in the *intensity* with which they hold those preferences’ (Rabinowitz and MacDonald, 1989: 93, emphases in original).

According to the directional conception of issue-based evaluations of parties by voters, the ‘neutral’ point or ‘status quo’ point in a given dimension of competition is of crucial significance; the utility that a voter associates with a party’s position is determined principally by whether or not the party is on the same side as the voter. If the party and the voter are both on the same side of an issue, relative to the status quo, then the voter’s evaluation of that party (on that issues) will be positive. If they are on opposite sides, then the evaluation will be negative. Under directional theory, increases in the extremism of the policy stance (or increases in distance from the ‘neutral’ point) are seen as indicators of the intensity with which a directional preference is held by a party. The intensity with which a preference is held by a party gives voters an indication of the likelihood that a policy will actually be enacted, should that party win power. Therefore, as the extremism of a party on the voter’s side of an issue point increases, so should the voter’s preference for that party. The strongest positive evaluations in directional theory occur when both voters *and* parties adopt extreme stances on the same side of an issue. For parties that are on the opposite side of the neutral point from the voter, the same logic means that parties on the ‘other’ side will be evaluated negatively, and increases in extremism on the ‘other’ side correspond to increasingly negative evaluations by voters.

Directional theory also provides that there is a ‘region of acceptability’ in politics, and that extremism beyond this region is punished by voters. Therefore, even if a party is ostensibly on a voter’s ‘side’ of the debate, more extremism on that side is only better up to a point. Sadly for VAA designers, the point at which the ‘region of acceptability’ is drawn is not theoretically specified, and empirical analyses often simply elide this issue by ignoring voter evaluations of ‘extreme’ parties and candidates. Westholm (1997) argues that the notions of a ‘region of acceptability’ and the non-specification of the scale of ‘penalties’ attributed by voters to parties falling outside of that zone lessen the falsifiability of the directional

approach, as these freely varying parameters can be adjusted to explain away findings that are not congruent with directional theory.

From the point of view of designing a MAUD rule, assuming as single dimension, the two theories result in divergent functional forms for issuing voting advice. Proximity theory indicates that as the absolute distance between a voter's position and a party's position decreases, the policy-based utility a voter associates with a party increases.

Or, formally that:

U_o increases as $|V - P_o|$ decreases. (1)

Where:

o is the o^{th} party from the set of existing parties;

U_o is the policy-based utility that a voter derives from the position of the o^{th} party;

V is the voter's position;

P_o is the o^{th} party's position.

This means that a MAUD rule built on proximity logic will rank the party with the least distance as 'closest' to the voter. Directional theory, on the other hand, indicates that the policy-based utility that a voter associates with a party increases as the *product* of voter and party positions increases (subject to the issue space being scaled around a '0' neutral or status quo point¹). So directional theory ranks the party with the highest product score as 'closest' or 'most in agreement' with the voter.

¹ Directional theory requires a scaling along the following lines. The neutral point, or status quo, is scaled to zero and the positions on either side of it are oppositely signed. Positions further from the centre are assigned

Formally, according to directional theory (and using the same notation as formula 1):

U_0 increases as $(V)(P_o)$ increases. (2)

Empirical research that has sought to test the relative merits of the two theories as predictors of voters' evaluations of parties offers little in the way of guidance to VAA designers. A series of articles has sought to demonstrate the superiority of one model over the other, indeed, in some instances; each side 'demonstrated' the superiority of their own model using identical data. Lewis and King (1999), surveying these investigations, came to the following conclusion:

'the empirical tests (...) amount to theoretical debates about which statistical assumption is right. The key statistical assumptions have not been empirically tested and, indeed, turn out to be effectively untestable with existing methods and data' (p.21).

Recent experimental evidence has sought to overcome some of these methodological limitations, with Classen (2009) indicating that such evidence points to spatial logic varying over policy areas. According to Classen's research, military spending-based evaluations of candidate positions conformed to a logic of proximity, whereas abortion-based evaluations conformed to a directional logic. In a similar vein, Cho and Endersby (2003) concluded, based on an examination of general election surveys in the UK that the proximity model performs better for understanding voters' evaluations of incumbent parties, while opposition parties' policy stances were better explained by a directional model. This line of research is potentially highly promising for VAA designers, it points to the possibility of

higher values, until the region of acceptability is breached, after which an unspecified punishment is doled out to parties.

multiple MAUD rules; depending on the issues considered and the parties that are being evaluated. In this research, however, in order to have a manageable and meaningful analysis, I restrict the analysis to a dichotomy of MAUD rules derived from proximity theory versus MAUD rules derived from directional theory. In constructing a directional MAUD rule, one does not specify *a priori* a value for the 'region of acceptability', because the theory gives no clear indication as to how such a value should be set.

From Issues to Dimensions

the notion of dimensionality, and the production of analyses that seek to reduce the complexities of political contestation to positions on a small subset of underlying dimensions is among the most potent weapons in the arsenal empirically oriented political scientists. The use of dimensions or cleavages of competitions to understand and compare the fundamental conflicts that shape political competition has a long and storied history, dating back to the incredible political transformations that occurred in late 18th century France. The seating positions of members of the Constituent Assembly that convened to write a new constitution were observed to correspond to the opinions, beliefs, and actions of the members (Asher and Shamir, 1983; Benoit and Laver, 2006). More recently, Lipsett and Rokkan's (1967) cleavage-based analysis of Western European party systems remains among the most influential accounts in political science, and empirical estimates of party positions on 'left-right' and other dimensions have been gathered over a wide range of countries using an array of techniques: analysing party manifestos by human coders (Budge *et al.*, 2001) and automated computer coding programs (Benoit *et al.*, 2009; Lowe *et al.*, 2011), surveying party elites (Farrell *et al.*, 2006), and the general public (Powell, 1999), monitoring the voting behaviour of party members (Hix, 2001; Jackman and Rivers, 2004), and conducting surveys of political experts (Castles and Mair 1984; Laver and Hunt 1992; Huber and Inglehart 1995; McElroy and Benoit, 2009; Marks and Steenbergen 2004).

In essence, political dimensions are a data reduction tool – they collapse party positions on hundreds, or even thousands of individual issues into a single indicator. Indeed, data

reduction techniques such as factor analysis and principal components analysis have been used in quantitative efforts to unpick the dimensions that underlie issue competition in several democracies (Benoit and Laver, 2006; Gabel and Huber, 2000; Mair 1989). The benefit of this approach is the increased simplicity and analytical tractability of the picture that emerges, relative to the garbled messages of real-world politics. However, as with any reductive approach, constraining political competition to fit onto a single dimension can result in a substantial loss of important information. Potentially, every additional political dimension considered in constructing an issue space adds to the depth of our understanding of the points at issue that separate political parties, and their supporters. However, there is a trade-off between detail and parsimony. As Benoit and Laver (2006) put it:

‘When we set sail across the Atlantic, for example, we would get lost if our only charts were so detailed that they show the position of every single grain of sand on every single beach we might pass. We need a description of the (political) world rich enough for the purpose at hand, but not so rich we cannot see the beach for the grains of sand’. (p. 18).

So, while the addition of dimensions may improve our ability to distinguish the nuances of party policy, it also complicates our conceptualisation of the political world. This *caveat* is especially worth bearing in mind for VAA site designers. Much of the appeal of such sites lies in their construction of a ‘political map’ which allows users to navigate the political world on the basis of a comparison of parties’ issue stances with their own.

The curvature of the political space

The third issue explored in this article is the manner in which distances or utilities should be aggregated over multiple dimensions. Assuming that the policy-based differences that separate parties and voters in a given political system require more than one dimension to be mapped, and given the task of a VAA designer, who seeks to provide users with an

indication of how which party they should vote for, on the basis of an 'overall' comparison of user preferences and party positions on several dimensions; it is important to bear in mind Benoit and Laver's (2006) observation that mathematicians have long noted that there exist an infinite number of possible geometries according to which distance can be calculated in a multi-dimensional space. The three-dimensional Euclidean space, *i.e.*, the functional geometry that most people use to view and interact with the world around them, is only one such geometry.

The intuitive appeal of Euclidean geometry, given its correspondence to our experience of reality, is perhaps unsurprising. Formal theorists in political science, working in the proximity tradition, have overwhelmingly favoured Euclidean geometry in modelling processes and predicting outcomes in politics. However, there is little evidence that this choice is based on an underlying body of evidence that describes how citizens and politicians actually make decisions. Benoit and Laver (2006) report a (2005) expert survey carried out by Laver which sought to uncover the underlying rationale for the popularity of Euclidean geometry among formal modellers. As with the debate on directional versus proximity theory, the results of this survey are rather dismaying for aspiring VAA designers. In explaining their use of Euclidean geometry, a vast majority of formal modellers stated that they employ Euclidean geometry because it produces models that are tractable using modern analytical techniques, with only a tiny minority (2 out of 23) claiming that their use of Euclidean geometry is related to an understanding of voters' decision-making. Indeed, the calculation of 'distance' in a two-dimensional plane according to Euclidean geometry is highly complex; it involves squaring the distance on each dimension, summing these squared distances, and then finding the square root of the sum of squared distance. An alternative geometry, the 'City Block' or 'Taxi Cab' metric simply provides that the distance on each dimension is summed by the voter. Westholm (1997) prefers the city block metric precisely on these grounds, arguing that it is more intuitive as a description of voter cognition than the 'complex logic' (p. 872) of information processing embodied in Euclidean geometry.

A broad variety of multi-dimensional geometries for calculating distances is incorporated in the Minkowski metric, which can be expressed as follows:

$$U_o = [\sum_j U_{oj}^r]^{1/r} \quad (3)$$

Where:

U_{oj} is the policy-based utility that a voter derives from the position of the o^{th} party;

o is the o^{th} party from the set of existing parties;

j the j^{th} issue from the set of issues;

U_{oj} is the policy-based utility that a voter derives from the position of the o^{th} party on the j^{th} issue;

and r is a real number.

The varying parameter that separates different metrics using formula 3 is 'r'. For Euclidean distances, r is set to 2. For city block distances, r is set to 1. Using the Minkowski metric, r can also be set to values lower than 1. The larger the value of the Minkowski metric r , the less 'compensatory' the model is. Values of r larger than 1 imply that a party will be preferred in the absence of large distances on any single issue, values of r smaller than 1 imply that a party will be preferred in the presence of a very small distance on a single issue. As such, values of r greater than 1 assume a 'risk averse' voter, who will heavily punish any large divergence. On the other hand, values of less than 1 imply a 'risk forgiving' voter, who will discount large differences and reward close proximity. Generally, this issue has received comparatively scant attention in the existing literature, though Benoit and Laver (2006) point to calculations of difference in computer science and to estimates of perceptions of difference and similarity in cognitive psychology as offering potentially interesting insights. Nonetheless, there is little research on which to build pre-conceptions with regards to the impact of the choice of Minkowski metric on the predictions that a VAA will generate. The

two most popular positions in the current literature are the 'Euclidean' ($r=2$) and 'city block' ($r=1$) geometries (Westholm, 1997). In the analysis presented in this article, we therefore examine the impact of varying r over two values $r = 1, r = 2$.

An uncertain field

This review of the literature on the appropriate type of MAUD for comparing voters and parties indicates that there exist multiple competing models and options in the political science literature, with no single method emerging as uniquely authoritative. In my research on the effect of MAUD selection under key question 2, I sought to examine the practical consequences of this uncertainty for a very specific applied project: namely the development of VEA websites which give voters feedback on their overall similarity to the policy platforms of political actors. Thus this research represents a substantive contribution to the broad literature on voter-party comparison, as well as representing a dramatic step forward in the research on the 'reliability' of VEA advice.

Beyond decision-rules

The popularity and impact of VEA projects in national elections across the world necessitates both transparency and a continuous effort towards site improvement among VEA-developers. To this end, it behoves VEA-developers and other scholars to investigate the quality, reliability and the validity of the outputs that VEA sites provide to their visitors (Gemenis 2011, p.2) from a wide range of perspectives, examining VEA development step-by-step to . Thus, a second aspect of my research in addressing key question 2 sought to facilitate such investigations by describing in detail the process of political mapping that has been developed by the 'Kieskompas' (Vote Compass) VEA developers in order to compare users' and parties' issue positions and their aggregation into deeper lying dimensions of party competition. Such detailed accounts of the various stages of creating a VEA are largely absent in the extant literature, and this account thus provides a guide to 'best practice', based on in-depth experience and a series of ideas for the improvement of VEA sites as contributors to the electoral process. Furthermore, describing the VEA-construction process

and the choices and trade-offs involved in that process, increase the overall transparency of VEA projects both for users and for members of the academic community.

- (key question 3) What can data gleaned from Kieskompas email surveys tell us about online survey response rates?

Survey response and response quality maximization are core issues of public opinion research, and the attention being devoted to these issues by practitioners is on the rise (Biemer, 2010; Curtin et al., 2000; Platek and Särndal, 2001). Online recruitment and collection provides potentially enormous benefits in terms of cost and speed of data entry. There also appear to be advantages to be exploited in terms of data quality: Chang and Krosnick's (2009) research indicates that, in terms response quality (as measured by random error, satisficing, and social desirability), Internet research outperforms random digit dialing. Within Internet research volunteer (i.e. non-probability opt in) samples had the highest quality responses.

However, while the quality of data among those who respond to surveys conducted online can be high, issues of non-response and bias are even more pronounced for online than for offline research, and methodologies for maximizing response rates and quality in an online-only environment are in their infancy. Furthermore, weighting procedures alone have proven insufficient to redress instances of bias using non-probability recruitment methods (Loosveldt and Sonck, 2008; see also Couper, 2000; 2011 for an overall typology of web-based surveys). Sills and Song (2002) conclude that 'low response rates, self-selectivity of Internet users, technological issues with the deployment of the research tool, and concerns over Internet security have troubled recent studies. Yet, for special populations that regularly use the Internet in their daily lives, the new medium has been found to be a sensible means of achieving meaningful results' (Sills and Song, 2002: 23).

As such, Internet surveying as a technique – or a survey mode – holds great potential for advancing our capacity to analyze public opinion (Dillman, 2007), with the combination of probability sampling and Internet surveying representing a particularly promising avenue

(Chang and Krosnick, 2009). However, assuring high levels of survey response is of particular concern for Internet surveys. We seek to investigate here whether the contents of the email soliciting survey response, a cost-free design feature of an online survey, can be manipulated to alter response rates and response quality.

While Internet surveying makes it easy to contact people in larger numbers and minimize sampling error, non-random 'opt-in' online panels face problems of significant coverage error, high response error relative to other modes, and potentially, measurement error (Couper, 2000; 2011), although Chang and Krosnick (2009) find that online survey modes perform well in terms of measurement. As such, to fully capitalize on the advantages offered by online public opinion data collection, it is important that academics and practitioners be intent on maximizing response rates and response quality in every aspect of the design of online surveys. The particular aspect of online survey implementation that I focused on in addressing key question 3 is the cover letter that accompanies the email request that is sent to potential respondents.

Historically, there is little evidence that content of survey cover letters has a strong effect on response rates in off-line survey environments (Harvey, 1987) although some research has pointed to small, but significant effects due to varying letter content (Brennan, 1992; Redline et al., 2004). Biemer's (2010: 817) discussion of 'Total Survey Error' as 'the accumulation of all errors that may arise in the design, collection, processing, and analysis of survey data' indicates that survey quality should be thought of in holistic terms, as the result of a multi-stage process of data generation. This approach has deep roots in the literature going back to Dillman's (1978) 'Total Design Method' (TDM), which focuses on maximizing survey response and quality at every stage of the survey design and implementation. Built on the premises of social response theory, the TDM seeks to orientate survey implementation and design process towards inclining the voter to perceive that the social and tangible rewards of participating in the survey outweigh the losses in terms of time and effort.

To be sure, the design and content of the cover letter that accompanies self-administered surveys is one of a range of elements to be considered in an overall survey design, and this

research should be seen as only one of many necessary contributions to improving response rate and answer quality. Several studies have found that techniques such as monetary incentives, advance (snail-mail) letters and telephone follow ups, questionnaire design, can serve to improve response rates and quality (Dillman, 2007; Christian et al., 2008; Rao et al., 2010). However, gains from these methods come with associated costs in terms of expense and time for public opinion researchers. Altering the content of the cover email, on the other hand, is virtually free of charge to practitioners and therefore should be an early step in designing any online survey.

However, a major difference between online and other survey formats is that digitized email messages take over the role of classic cover letters or of human interviewer contacts. Such email 'cover letter' contains the link where the potential respondent can access the online survey. Thus, the content of email messages may more important for online surveys than cover letters were for paper questionnaires. The research carried out to investigate key question 3 thus contributes to the extant literature principally by providing objective evidence of the scale of the effects of choices concerning cover letter content on response rates and answer quality for a non-probability online panel survey. It also seeks to draw practical lessons for practitioners on how best to tailor their cover emails to maximize response and quality on the basis of this evidence.

3. Measurement, Data and Methods

- Key Question 1: Are users of VAA/VEA websites influenced by the ‘advice’ that they receive online? Specifically, are they more likely to vote for a party that they have been ‘advised’ to vote for online?

The research design that I developed to address this question was as follows: each user of kieskompas.nl between the site launch date (April 27th, 2010) and the day of the election (June 9th, 2010) created a ‘log file’, which is a row in a data matrix that keeps a record of the details of their visit to the site. Immediately before they received their ‘recommendation’ (i.e., the party that was deemed ‘closest’ to them by the VAA), users were asked to estimate the likelihood that they will ever vote for each of the main parties competing in the election. Eijk et al. (2006) argue that this survey item is a useful way to operationalize the electoral utility that a voter attaches to each party competing for their votes at the election, and that it captures richer information on a respondent’s electoral preferences than a dichotomous vote intention item. Responses to this item were constrained to natural numbers ranging from 0 to 10, where 0 means ‘not at all probable’ and 10 means ‘very probable’ (all 11 parties were evaluated on this scale by the respondent). Users’ stated probabilities to vote (hereafter, PTVs) for each party were recorded, along with the specific advice that they received, in their log files.

The second stage of data collection relied on the site’s ‘extra survey’ feature. This was an opt-in section of the site where users could register their email addresses and indicate their permission to be contacted for subsequent surveys. Of the over 800,000 unique visits generated by the site, just over 8,000 users left emails and indicated permission to be contacted. The email addresses of users were recorded in their log files, in the same row of data that contained the other details of their site visit (including their PTVs for all parties and the advice that they received). Working with Dr. Krouwel and an external agency (‘Synovate’ a global public opinion research company and partner of Kieskompas) I sent a

post-election survey to these users, which included an item on how they had voted, and items asking respondents for their perceptions of the influence that visiting kieskompas.nl had on their vote choices. The post-election survey also asked users to recall the advice that they received, which enabled me to analyse recall data for accuracy and bias as discussed in the previous section. The survey elicited a response rate of just over 52%, generating an *N* of 4,257 respondents.

Each survey request was marked with a unique identifier number, which was identical to the unique identifier number in the respondent's log file entry. Using Excel's 'VLOOKUP' command, we were able to use these unique identifiers to merge users' responses to the post-election survey with their individual log file entries. This approach meant that we captured the three vital elements required for constructing a plausible individual-level analysis of the effects of a VAA on user vote choice: 1) each user's pre-advice party electoral utilities, 2) an objective measure of the advice received by each user, and 3) recall data on the vote choice of each user. Data analysis was performed on STATA 11 and SPSS 18 statistical packages, and I detail the results and analysis in the next section.

- (Key question 2) How does the design and execution of VEA sites influence their contribution to both the accountability and the legitimacy of elections? How can VEA site design be improved to maximise this contribution?

In this section, I will focus mostly on my research on the effects of MAUD rules for the output of VEA sites. In order to assess the impact of choosing any of alternative MAUD rules on the output that a VAA site would generate, I used the log files of over 45,000 users of the Kieskompas.nl website in the 2010 Dutch elections to simulate the advice that the site would have produced using each rule. While the site's logfiles recorded over 800,000 unique visits, I selected only the logfiles of those users who completed all 30 issue questions, rated their likelihood to vote for each of the main parties, and responded to at least some of the 'extra questions' asked of users after they receive their advice. I did so in order to minimize

the contribution of repeat users and spurious entries to our analysis. In order to perform the simulations, I re-ran the vote advice process for all 45,000 log files using each of the 6 following MAUD rules.

1. Proximity based, single dimension. The winning party is the one that has the smallest absolute distance from the user on the 'left-right' x axis.
2. Proximity-based, 2 Dimensions, Euclidean metric. The winning party is the one with the smallest Euclidean distance from the user on the 'left-right' and progressive-conservative' axes. This was the approach used by the kieskompas.nl site.
3. Proximity-based, 2 Dimensions, city-block metric. The winning party is the one with the smallest summed distance from the users on the on the 'left-right' and progressive-conservative' axes.
4. Proximity-based, 30 Dimensions, Euclidean metric. The winning party is the one with the smallest Euclidean distance from the user over all 30 political issues measured by kieskompas.nl.
5. Proximity-based, 30 Dimensions, city-block metric. The winning party is the one with the smallest summed distance from the users over all 30 political issues measured by kieskompas.nl.
6. Directional-based, 2 Dimensions. The winning party is the one with the largest score when the product of party and voter positions on the x and y axes is summed.
7. Directional-based, 30 Dimensions. The winning party is the one with the largest score when the product of party and voter positions over all 30 issues is summed.

For two-dimensional MAUD rules, I used the x and y axis co-ordinates of users and parties generated by averaging their positions on 'left-right' and 'progressive-conservative' issues, respectively. For thirty-dimensional MAUD rules, I compared the responses of users and parties on each issue separately. Having thus generated 6 separate sets of site outputs, the

analysis (detailed in the subsequent section of this report) proceeded to examine how these compare by examining the aggregate ‘volatility’ engendered by changing MAUD rules, investigating how much of an overall difference changing from one rule to another represents in terms of aggregate site output. Secondly, I analysed whether certain parties, and, indeed, certain types of parties, benefit or suffer in terms of the advice that is issued, depending on the type of MAUD rules employed.

In assessing the broader questions regarding VEA execution I relied primarily on qualitative data derived from my experience working on VEA websites during my time at Kieskompas. Specifically, I worked from the beginning to the end of two major projects: sites in Canada and Turkey. In piecing together a detailed account of the ‘Kieskompas method’ I collaborated extensively with Dr. Krouwel, reflecting on and codifying the procedures that Kieskompas has developed in developing VEA websites, and analysing possible alternative approaches that could have been taken at each stage.

- (Key question 3) What can data gleaned from Kieskompas email surveys tell us about online survey response rates?

The experimental study conducted to investigate key question 3 focused on three aspects of the email message that online survey respondents receive: the type of appeal made to the respondent, the complexity of the message, in terms of writing style, and the tone of the email. We composed 8 email ‘cover letters’ which comprise all possible combinations of these dimensions. A large panel of Internet survey volunteers (around 11, 300 respondents) was then randomly assigned to one of 8 groups.

In our letter formulation process, Dr. Krouwel, Mr. Fazekas and I conceptualized the online ‘cover letter’ as a persuasive document, designed to influence the motivations of

respondents in such a way that they were more likely to respond to the survey. This approach builds on Dillman's (2007) contention that the interaction that takes place between a surveyor and a survey respondent is best conceptualized as a social exchange. Dillman's model of this exchange focuses on three dimensions: *reward*, *cost*, and *trust*. His formula for maximizing survey quality in implementation involves deploying measures which build trust (for instance, accompanying a survey request with a pre-paid sum, as well as via multiple contacts) and improve respondents' perception of the rewards that they will accrue from participating (such rewards can be either material or immaterial) while reducing the perceived costs of participating (in terms of time, effort, and, potentially, the risk of confidentiality breaches). Overall, Dillman's approach leads us to focus on the motivations of respondents, and to treat respondents as intelligent social beings – meaning that one must maximize those elements that will lead respondents to positively evaluate participating in the survey, and will motivate them to complete each survey diligently.

The study developed three dimensions of letter variance designed to capture the three-dimensional matrix of survey participation developed by Dillman (1978), which focus on the reward, cost, and trust elements of survey completion, as a form of societal interaction. Combining all three dimensions generates 8 letter types (2 x 2 x 2): altruistic-simple-informal (ASI), altruistic-simple-formal (ASF), altruistic-complex-informal (ACI), altruistic-complex-formal (ACF), egoistic-simple-informal (ESI), egoistic-simple-formal (ESF), egoistic-complex-informal (ECI), egoistic-complex-formal (ECF).

In terms of an online cover letter – we sought to find a method for manipulating each of these three dimensions. The first dimension that we considered was *reward*. Given that no

monetary incentives were offered in our design, it appears unlikely that a survey cover letter, in and of itself, can be manipulated to affect the material reward/cost structure of the survey request. However, Hansen's (1980) 'self perception' model of survey response indicates that survey requests can offer internal motivators to encourage survey response by associating survey completion with either personal or societal rewards. We managed this by weaving a specific appeal into the text of the letter. These appeals do not ask respondents to fill out the survey to maximize survey accuracy, but rather they emphasize the intangible personal benefits that come from expressing one's opinions, or the benefit that the research represents to society as a whole. The former appeals can be considered 'egoistic' (appealing to the respondents' sense of self) and the latter 'altruistic' (appealing to respondents' sense of social obligation) (Redline et al. 2004). We therefore divided letters into two types of underlying appeal: egoistic and altruistic. The 'egoistic' letters focus on the respondent as individual, emphasizing how important, reliable and valuable the respondents' opinion is. The altruistic letters, on the other hand, build on the idea of reciprocity, contribution to research and society being achieved through responding to the survey.

According to these premises, for egoistic messages we employed a vocabulary that emphasized words such as "you", "your", "opinion", whereas for altruistic messages words as "contribute", "contribution", "help". Of course, all of these framing efforts take place within the constraints that the email is comprehensible, communicates the topic of the survey, and asks politely for a response: therefore, for instance, the word "help" was not completely dropped from egoistic messages. Using text mining, the appearance and frequency of the previously mentioned set of words was analyzed. The words 'contribute'

and 'contribution' appear only in the altruistic messages (sparsity of 80% and 90%), whereas 'opinion' only appears in the egoistic messages (sparsity of 60%, only appearing in egoistic messages). The use of 'you' and 'your' is much higher for egoistic messages than altruistic messages. 'Help' appears much more often in altruistic messages, but it is also present for the egoistic messages, where it appears once. The low frequency counts for the major words on both dimensions reflect the functional requirement that the messages not to be overly long (no letter was longer than 250 words).

The second dimension of Dillman's (1978) social exchange matrix is *cost*. The cost of a survey completion is the time, effort and risk that a respondent perceives to be associated with completing a survey. In terms of an email cover letter, we argue that the linguistic complexity of the survey proposition can act as a proxy for the cost of completion. This is a particularly important element for online surveys, where the content and design of the survey itself is typically not contained in the contact email. All users have to go on in evaluating the survey cost is the email 'cover letter'. We therefore posited that complex language patterns in the cover letter may serve to increase the perceived cost of completion, and consequently depress response rates. Simple messages, on the other hand should minimize the cost perception of respondents, and improve response rates. We therefore divide email messages into simple and complex. Simple messages were written in words with small numbers of syllables, and avoiding multi-clause sentence structures. Complex formulations employed longer wording and complex sentence structures. Again, to maintain realism – the simple formulations were not childishly simple and the complex formulation was not impenetrably complex. As such, integrating this design element into letter composition involved a degree of subjective creativity, however we were able to

examine the extent to which 'complex' letters differ using a computer algorithm designed to distinguish textual complexity.

As a benchmark for the difficulty of the text, we employed the Flesch-Kincaid readability scores for each email. Here, the lower scores indicate easier understanding, or more precisely lower number of grades in formal education required for comprehension. The Flesch-Kincaid scores indicate that 'complex' messages are much harder to read. However, tone, which we shall discuss next, is not unrelated to complexity. The combination of a 'simple' formulation and an 'informal' tone generated the easiest messages to read (6.27, respectively 7.56), while 'formal' messages were more difficult to read across the board, meaning that the most complex formulations were the two combining formal tones and complex formulations.

Thirdly, in as much as one can, we sought to manipulate tone of the cover letters. Given the novelty of this research in a Dutch cultural context, it is difficult to predict the effects of letter tone. On the one hand, a formal tone designates legitimacy and authority. For instance, Brennan (1992) found that cover letters signed by researchers whose title designated high status on the research team generated higher completion rates than cover letters signed by low ranking researchers. On the other hand, a friendly tone can establish a positive and trusting communication stream. It seems most likely that respondents' interpretation of tone is a result of societal or personal factors – some people may place greater trust in formal letters others in informal letters.

In terms of writing the letters, the tonal dimension was the easiest of the three to manipulate, since the Dutch language has separate formulations for "You" (formal) and

“you” (informal), essentially covering this dimension perfectly. However, to make sure that the letters capture substantively different tones, we also altered the opening and closing sections of the letters. Whereas, in the formal versions we posit the topic of the survey and introduce Kieskompas, the informal messages start with ‘Hi’ or ‘Greetings’. For informal messages, we also modified the first sentence, so that the Kieskompas Director introduces himself by name to induce a more personal, informal, and closer atmosphere. Furthermore, in the closing lines we used “Kind regards” in the informal messages, instead of “Yours sincerely” in the formal messages – followed by signature in both cases.

A final step in our analysis of our operationalisations of the concepts of appeal, complexity and tone involves Wordscores (Laver et al. 2003). Before running the analyses, one must emphasize that the scores obtained here should be interpreted with caution for at least two reasons. First, the texts are relatively short, but more importantly, as we have described, there are three underlying dimensions in the texts, and Wordscores is suitable to detect one dimension, based on the reference texts chosen. However, we consider this approach informative because it offers a way to put uncertainty measures around our treatment differentiation. As both the complexity and the tone are relatively straightforward in assessment, we focus the Wordscores analysis on the altruistic vs. egoistic dimension.

ACF and ECF were chosen as reference texts for this analysis, because ECF is the longest and richest text, and the counterpart on the altruistic end of the scale was needed, to have two similar anchors (even though it is not the longest text). After running the analysis on the raw texts – with punctuation removed, Yet again, the low number of words for each email and the fact that only 68 % of the words were scoreable makes us treat the results with caution.

The point estimates on average look promising, and they essentially suggest that the text manipulations were successful. All the letters employing an egoistic appeal were found to be substantively and statistically different from the altruistic emails. Egoistic messages crowd around the lower end of the scale, whereas the altruistic ones are estimated to be around the upper end of the scale. Based on our reference text designation (ECF low, ACF high), these results were as expected and provided confirmation that our treatment design was capturing the relevant theoretical constructs.

I now elaborate here on how the experiment was designed and implemented. Our design uses a panel of respondents who left an email address and indicated consent for being occasionally re-contacted after completing a Dutch online vote advice application:

Kieskompas.nl. Several members of this panel have been already contacted for political surveys designed by Kieskompas and implemented by the public opinion agency Synovate.

This approach suggests a possible limitation of our study related to self-selection.

Accordingly, individuals who were interested enough to previously express their opinions and further help social scientific research make up our full database. This clearly shows that they are, on average, a politically more interested, motivated, and active segment of the general population. However, our goal is not to generalize our findings to the level of Dutch people. Moreover, this aspect of our sample also works against us in a rather peculiar way: it suggests that we should **not** anticipate that email content has a large impact on response rates or response quality, since members of the panel have already had experience with surveys and they are obviously interested in the political arena. Nevertheless, as we will see below, there is evidence that letter contents had significant effects on our indicators of concern.

The original (pooled) dataset had 11,374 respondents. These respondents can be grouped into three categories that are relevant for our study: (1) 4,901 respondents who had not yet been re-contacted by Kieskompas and Synovate, (2) 1,812 respondents who had been re-contacted once already, and (3) 4,661 respondents who had been re-contacted twice already. This categorization is important because depending on the frequency of previous collaboration we may expect different response behavior. The two previous survey waves had elicited the following response rates: the first wave (sent to respondents from category 3 above) had a 62% response rate (3,619/5824), while the second wave (this was the dataset used in the study of key question 1) elicited at 52% response rate (4,257/8146). Both waves were implemented very close to the panel sign-up that took place during the Dutch election campaigns in May and June 2010. The topics studied in these waves were leadership evaluations, media consumption, vote choice, and perceptions of the Kieskompas.nl product. The 4901 respondents who had not been re-contacted at the start of the experiment had left their emails on the Kieskompas.nl local elections websites in March 2010.

The resultant data collection strategy led us to designate 8 groups and we used block randomization to assign individuals to these groups. We chose this method because we wanted to minimize the between-group variation for three dimensions: gender, education, and previous contacts. From a theoretical perspective, all three dimensions are expected to contribute substantively to how these letters are perceived and what behavior do they trigger. Having an experimental group that is well above average in number of contacts or women are over-represented would lead to artificial between-group differences in our

measures of response rates and quality. After the block randomization of the groups, a letter was randomly assigned to each group.

The survey concerned coalition formation following the national elections, and was sent to respondents between the 12th and 22nd of December 2011. All letters indicated clearly that this survey is about the coalition formation and the Dutch government. It is possible that the topic had some deflating effect on response rate however, the considerable time that had elapsed between previous contact (either in signing up via the local election sites in March, or receiving a post election survey in June) probably explains why response rates generally were lower for this wave than previous ones. Our analysis (described in detail in the next section) focused on differences between the groups in terms of response rates and response quality – response quality was measured using items that were reversed – in this case we are focused on the consistency of responses when framing/wording is varied.

4. Results and Analysis

- Key Question 1: Are users of VAA/VEA websites influenced by the ‘advice’ that they receive online? Specifically, are they more likely to vote for a party that they have been ‘advised’ to vote for online?

I argue in my analysis of the data collected to address key question 1 that VAA sites are, fundamentally, heuristic-generating instruments. The heuristics that they generate for their users are most analogous to endorsements (with each individual given a specific party endorsement by the site), but they also have an element of ideological heuristic content (because voters are given a simplifying indication of their overall issue-based similarity to the competing parties). The fact that the output of the site is presented as reflecting the user’s own opinions, rather than the opinions of any particular group or party, gives the site endorsement/recommendation a powerful persuasive element. Furthermore, as demonstrated in Figure 1 (where the ‘advice’ of the VEA site used in this investigation is circled in red), the advice is typically easy to understand: the ‘closest’ party, according to the VAA measurement procedure, is clearly identified for each user. Thus a VAA’s advice provides a sense of personalisation and neutrality combined with clear presentation – making it potentially very persuasive. This line of reasoning leads us to anticipate that at least some voters will be swayed by the vote recommendation provided by their visit to a VAA site.

Figure 1. Advice screen of Kieskompas.nl 2010 Dutch legislative elections (closest-party advice circled).



Specifically, I test the following hypothesis: H_1 : The effects of VAA site advice on users should lead them to be more likely to vote for parties recommended by the VAA site than for non-recommended parties, *ceteris paribus*.

However, the classic 'minimal effects' counter-argument to the contention that campaign information can change voters' preferences holds that people do not absorb political information in an undifferentiated manner, nor do they easily change their partisan preferences on the basis of new information. Rather, a rich tradition in public opinion and voting behavior research going back to the work of Berelson, Lazarsfeld, and McFEE (1954) has argued that political information is differentially processed according to each individual's partisan and value-based predispositions, meaning that individuals tend to absorb information that is congruent with their predispositions more frequently and readily than they absorb ideologically or politically incongruent information (Zaller, 1992). Iyengar and Simon (2000: 158-161) contend that the traditional 'hypodermic' model of campaign effects fails to capture this aspect of voter cognition, and they argue that the more recent 'resonance model', which views campaign messages as interacting with the prior

predispositions of voters, is a more realistic view of the complexities of campaign effects. One of the key insights of the resonance model is that ‘messages that are counter-attitudinal will be actively resisted, whereas that are consonant will be accepted’ (*ibid.*: 159).

We would therefore be surprised if VAA users uncritically acted on the advice given to them online by an automated algorithm, no matter how much that advice simplifies political decision making, if that advice runs radically contrary to their prior political predispositions. VAA advice is just one of a number of competing information sources to which a voter may be exposed during a campaign. When such cues provide conflicting advice, it is the voter herself who must decide which information to accept (Lupia and McCubbins; 1998: 40). Zaller’s (1992) Resistance Axiom, the insights of the communications, political psychology and electoral behaviour literatures, and, indeed, common sense would lead us to anticipate that users will resist integrating highly incongruent information into their political considerations far more than they will resist information that chimes with their existing preferences. As such, it is unlikely that users will vote for a party that they had previously completely ruled out, just because of the advice of a VAA site.

I therefore argue that the most likely political effect of a VAA recommendation should be to increase one’s inclination to vote for the recommended party, *when one is already seriously contemplating voting for that party*. When the recommended party has been deemed electorally off-limits by the user, we do not expect the advice to influence their vote choice. As such, the impact of the information received by users on VAA sites may be largely confirmatory, rather than persuasive. This line of reasoning leads to a second hypothesis:

H₂: VAA recommendations will be most influential when the recommended party is already being seriously considered as a potential vote choice by the user, and will have little or no influence when the recommended party is not being considered as a potential vote choice by the user.

Descriptive analysis

I begin by looking at the evaluations made by users themselves of the influence that their visit to kieskompas.nl exerted on their vote choices. H_1 would lead us to anticipate that a large portion of users should consider that the site exerted some sort of influence on their vote choice, while H_2 suggests that this influence should primarily be one of preference confirmation (rather than preference conversion). H_2 also implies that those who received advice that was congruent with their partisan predispositions should attribute greater influence to their visit to the site than those who received incongruent advice.

Table 1 presents the response frequencies for a categorical survey item, which asked users to characterise what, if any, effect kieskompas.nl had on their vote choice. We can see that, while the second largest group, at 29%, stated that their visit had 'no effect', 71% of users choose options indicating that they consciously experienced some form of VAA 'effect'. Just over 30% indicated that the experience was one of preference confirmation, while 26.1% stated that their visit presented them with previously unconsidered options. A smaller number indicated that the visit directly shaped their vote choice – either by helping them to choose among several parties that they were considering (9.2%) or by leading them to vote for a previously unconsidered party (1.2%).

Table 1. Responses of survey group to the question: ‘Which of the following best describes the influence that visiting kieskompas.nl had on your vote choice?’

Type of Influence	Frequency	Percent
Confirmed my preferences	1,195	30.2
No Effect	1,145	29.0
Showed me other options, but didn't change my mind	1,032	26.1
Helped to choose among several parties	365	9.2
Lead me to vote for a party I had not previously considered	45	1.1
Other	171	4.3
Total	3,953	100

Users’ perceptions of the strength of the kieskompas.nl effect was also measured on a 0-10 scale, where 0 is ‘no effect’ and 10 is ‘a very strong effect’. Overall, ‘0’, which was the response of 35.3% of users, is the largest category in terms of frequency. Nonetheless, a clear majority of respondents estimated that the site had influenced their vote to some extent – and, among users who estimated at least some effect, 5, 6, and 7 were the most common estimates of the strength of that influence. Very few voters attributed the highest level of influence to their visit to the site: less than 1% scored the influence of their visit as a ‘10’.

Inferential Analysis

A logical starting point for an inferential analysis of key question 1 is to look at how many respondents voted for the party that kieskompas.nl recommended to them. In the discussion that follows, for the sake of simplicity, I adopt a reductive nomenclature. Those respondents who reported voting for the party that kieskompas.nl recommended to them are described as ‘followers’. The term ‘non followers’ designates respondents who voted for any party other than the one recommended to them. Among all survey respondents, 26.5% were followers. The obvious implication is that 73.5% were non followers, and so advice ‘following’ was not a majority phenomenon among the users surveyed. Of course, raw numbers of followers versus non followers tell us little about whether the site exercised any discernible influence of user vote choice. After all, the site is designed to show users how closely their policy preferences align with those of the parties competing in the election, and a huge volume of work in political science (spanning back to Downs, 1957) tells us that ideological closeness is a major driver of vote choice. We therefore cannot say that 26.5% of the respondents ‘followed’ the advice that they received online. Taken in isolation, these figures could simply mean that the site correctly identified the vote intentions of 26.5% of its users on the basis of their policy positions.

The crucial element for understanding the nature of the influence exerted by the kieskompas.nl site on its users lays the interplay between users’ pre-existing preferences and the advice that the site generates. A simple cross tabulation, provided in table 2, tells much of the story that we develop in more detail in this analysis. Table 2 describes the percentages of ‘followers’ and ‘non-followers’ of kieskompas.nl advice for two groups. Table 2 divides respondents according to the relationship between the advice given and each user’s pre-advice political preferences, separating those who received ‘congruent’ voting advice (i.e., advice to vote for a party to whom they were given a PTV of 6 or higher) from those who did not. Table 2 shows that people were only likely to ‘follow’ the advice issued to them when the recommended party was one that they were strongly considering voting for already. Only 157 survey respondents (just under 4% of the entire sample) followed the

site’s advice when they had estimated their pre-advice likelihood of ever voting for the advised party at 5 or lower on the 0-10 PTV scale.

Table 2. Was VAA advice ‘followed’? Congruent versus non-congruent advice (column percentages in parentheses).

Did the respondent vote for the party recommended by kieskompas.nl?	Did the respondent rate their likelihood to vote for the advised party at 6 or higher out of 10?		
	No	Yes	Total
No	2,456 (94.0%)	531 (36.7%)	2,987 (73.5%)
Yes	157 (6.0%)	918 (63.3%)	1,075 (26.5%)
Total			

The next step in this analysis required that I restructure the data. For this part of our analysis, the dataset has been ‘stacked’ on the Vote, Advice, and PTV variables. This means that we restructured the dataset so that each user-party relationship of interest is considered as an observation in its own right, occupying a row in the data structure. With 11 parties being considered, each individual respondent is therefore represented as 11 ‘observations’ in the data. The advantage of this approach is that it allows us to consider the relationships between variables such as ‘Advice’, ‘PTV’ and ‘Vote’ across all parties in a single analysis, rather than having to analyse each relationship over all 11 parties. In the conventional ‘wide’ dataset structure, which we used to produce tables 1 and 2 and Figure 2, each respondent is considered as a single observation, denoted by their occupation of a single row in the dataset with separate columns capturing the relationship between that

respondent and each party. See Eijk et al. (2006) for details on the stacking procedure. The stacking procedure leads to considering variables of interest over 40,524 observations, where each observation is a user-party dyad (i.e., $3,684^2$ respondents multiplied by 11 evaluated parties). This restructuring of the data allowed me to split voter-party dyads into two groups: 'recommended' dyads (where the party in the dyad was recommended to the user by kieskompas.nl) and 'not recommended' dyads. Each voter can thus be represented as 1 'recommended' voter-party dyad, and 10 'not recommended' voter party dyads. Overall, this means that we compare a group of 3,684 'recommended' voter-party dyads to a group of 36,840 'not recommended' voter-party dyads. Each dyad also has a value for 'PTV' - which indicates the opinion of the voter as to the likelihood that they will ever vote for the party in that dyad.

A first step in the analysis here looks at the distribution of PTV scores for voter-party dyads in the recommended and not recommended groups. These distributions are shown in table 3. On average, as one would anticipate, given that the kieskompas.nl VAA sought to match users with ideologically 'close' parties in its recommendations, the recommended dyads have a higher PTV than the not recommended dyads. The average value PTV value for recommended dyads is 5.3, while the average for not recommended dyads is 3.2, a difference which is statistically significant over our 40,524 observations with 99% confidence using a t-test. This finding means that we must hold PTV constant to isolate any causal effects that may be attributable to the VAA advice. However, as we can see in table 3, there are a large number of instances of users being advised to vote for parties to which they had given very low PTV values, meaning that there are large enough numbers to compare 'recommended' and 'not recommended' dyads at all 11 values of PTV.

² Respondents who did not rate one of the parties on the PTV survey item, who did not vote, who did not recall their vote, or who reported a vote for 'Other' were dropped from this part of our analysis.

Table 3. Distributions of 'PTV' for recommended versus not recommended voter-party dyads.

PTV	Recommended dyads	Not recommended Dyads
10	14.1% (521)	3.1% (1,140)
9	9.9% (364)	3.8% (1,399)
8	11.7% (432)	6.5% (2,383)
7	9.5% (351)	7.8% (2,877)
6	8.7% (320)	8.0% (2,957)
5	8.6% (316)	7.9% (2,906)
4	5.2% (192)	5.7% (2,105)
3	4.5% (167)	5.6% (2,067)
2	4.7% (174)	6.15 (2,253)
1	4.3% (157)	6.8% (2,512)
0	18.7% (690)	38.7% (14,241)
Total	100% (3,684)	100% (36,840)
Average PTV	5.3	3.2

The goal in collecting and transforming these data was to investigate the relationships between 'Vote', and 'Recommended', while holding 'PTV' constant. 'Vote' is thus the dependent variable, it is dichotomous and is coded '1' when a respondent reported voting

for the party, and '0' when they did not. 'Recommended' is also a dichotomous variable that is coded '1' when kieskompas.nl recommended the party in a voter-party dyad and '0' when it did not. 'PTV' is the respondent's self-reported pre-advice likelihood of ever voting for the party in the user-party dyad.

The empirical implications of H_1 for these data are that user-party dyads where 'Recommended' equals '1' should be more likely to be 'Voted' than dyads where 'Recommended' equals '0', when PTV is held constant. H_2 leads us to anticipate that the differences between the 'recommended' and 'not recommended' dyads should be most pronounced where high PTV values indicate that the advised party was being seriously considered during the election campaign, and less pronounced among dyads with small PTV values.

In order to test whether these empirical implications are borne out, I segmented the data according to the PTV of the user-party dyads. Table 4 shows the results of this analysis. This computation allows for a comparison of 'recommended' versus 'not recommended' user-party dyads separately at each value of PTV. I was thus able to test whether the proportion of 'voted' user-party dyads is higher for 'recommended' dyads than for 'not recommended' dyads, holding constant the pre-advice likelihood of the respondent to vote for the party. For each value of PTV in table 4, I report a Pearson's chi-squared statistic and associated p-values with 1 degree of freedom for the proposition that the proportions of 'voted' and 'not voted' dyads are identical for the groups of 'recommended' and 'not recommended' dyads. We take p-values of .05 or smaller as a cut-off point for considering the difference in proportions across groups statistically significant.

We can see from table 4 that recommended parties are significantly more likely to become voted parties than not-recommended parties (confirming H_1), and this effect is principally observed at PTVs of 6 or higher (confirming H_2). The scale of the effect is largest at PTV values of 8, 9, and 10, where recommended user-party dyads are 'voted' dyads 18.2%-22%

more often than not recommended dyads. The VAA effect is smaller, but still statistically significant at PTVs of 6 and 7, where recommended dyads are voted 7.7% and 8.5% more often than not recommended dyads, respectively. For values of PTV lower than 6, only one category (3) shows any significant difference across groups, and the difference here is small, at 2.6%. For other low PTV values, there is no discernible difference between the recommended and not-recommended dyads in terms of their likelihood to be ‘voted’ dyads, meaning that we cannot identify any VAA ‘effect’ for these values of PTV.

Table 4. Proportions of ‘voted’ user-party dyads for ‘recommended’ and ‘not recommended’ dyads for each value of PTV (N for each proportion in parentheses, statistically significant differences in boldface).

PTV	% ‘Voted’ dyads Recommended	% ‘Voted’ dyads Not Recommended	% Difference	Chi-Squared test results
10	84.1% (438/521)	65.9% (751/1,140)	+18.2%	$\chi^2(1, N = 1,661) = 58.18$ $p = .00$
9	62.6% (228/364)	40.6% (568/1,399)	+22.0%	$\chi^2(1, N = 1,763) = 56.64$ $p = .00$
8	43.5% (188/432)	24.5% (584/2,383)	+19.0%	$\chi^2(1, N = 2,815) = 66.41$ $p = .00$
7	21.9% (77/351)	13.4% (385/2,877)	+8.5%	$\chi^2(1, N = 3,228) = 18.75,$ $p = .00$
6	13.1% (42/320)	5.4% (159/2,95)	+7.7%	$\chi^2(1, N = 3,277) = 30.13$ $p = .00$

5	4.7%	3.3%	+1.4%	$\chi^2(1, N = 3,222) = 1.78$ $p = .18$
	(15/316)	(96/2,906)		
4	1.6%	1.5%	+0.1%	$\chi^2(1, N = 2,297) = .00$ $p = .96$
	(3/192)	(32/2,105)		
3	3.6%	.9%	+2.7%	$\chi^2(1, N = 2,234) = 10.77$ $p = .001$
	(6/167)	(18/2,067)		
2	1.7%	1.2%	+0.5%	$\chi^2(1, N = 2,427) = .44,$ $p = .51$
	(3/174)	(26/2,253)		
1	.6%	.5%	+0.1%	$\chi^2(1, N = 2,669) = .04,$ $p = .84$
	(1/157)	(13/2,512)		
0	0%	.4%	-0.4%	$\chi^2(1, N = 14,931) = 2.47,$ $p = .12$
	(0/690)	(51/14,241)		

These findings accord with my characterisation of the effects of a VAA recommendations on user vote choice in H₁ and H₂. We can see that VAA users rarely have their prior preferences radically overturned due to getting a recommendation online, but it appears that a congruent recommendation can help to firm up a voter's mind when they are already leaning towards the recommended party.

- (Key question 2) How does the design and execution of VEA sites influence their contribution to both the accountability and the legitimacy of elections? How can VEA site design be improved to maximise this contribution?

I begin my summary of analysis with my findings on the effects of variation in MAUD rule, before going on to discuss the outlining of the key stages in the overall design of a VEA site. Overall, I restricted my analysis of the effects of MAUD rule selection to three sources of variation faced by VAA designers in choosing a MAUD rule (these sources of disagreement are discussed in detail in the literature review section of this report):

- 1) The assumption of 'proximity' versus 'directional' theory in deriving voter utilities from a comparison of their positions with party positions.
- 2) 1 versus 2 versus 30 dimensional issue-spaces.
- 3) Euclidean versus City Block distance metrics.

For MAUDs derived from directional theory, I simply summed the directional utilities on issues considered and recommend that the voter choose the party with the highest value. For 'proximity' issue spaces I compared Euclidean to city block metrics in aggregating distances over multiple dimensions, and the simulation recommends that users vote for the party who is 'closest' to them, according to each calculation. This approach leaves us with 7 discrete MAUD rules, 5 derived from proximity theory, and two derived from directional theory, which we consider in our simulation. These rules are summarised in Table 5.

Table 5. Types of MAUD rules explored in this report.

	1 Dimension	2 Dimensions	30 Dimensions
Proximity based:	1. Lowest absolute distance on left-right axis.	2. Lowest Euclidean distance (Kieskompas.nl method) 3. Lowest City Block distance	4. Lowest Euclidean distance 5. Lowest City Block distance
Directional	(not considered as awards all recommendations to most extreme parties).	6. Highest sum of directional utilities.	7. Highest sum of directional utilities.

In order to perform the simulations, I re-ran the vote advice process for all 45,515 log files using each of the MAUD rules described in Table 2. Having thus generated 7 separate sets of site outputs, I begin my analysis by examining the aggregate ‘volatility’ engendered by changing MAUD rules, in terms of the proportions of ‘first preference’ advices issued. Secondly, I analyse whether certain parties, and, indeed, certain types of parties, benefit or suffer in terms of the advice that is issued, depending on the type of MAUD rules employed.

Results

Table 6 records the aggregate output generated by the simulations using the competing MAUD rules described in the previous section. The parties which do particularly well in

terms of aggregate numbers of recommendations according to each MAUD are circled in red.

Table 6. Aggregate Partisan Recommendations generated by Alternative MAUD Rules (% of first choice recommendations per party)

Party	1. Prox. 1-D	2. Prox. 2D, Eucl.	3. Prox. 2D, CB	4. Prox. 30D, Eucl.	5. Prox. 30D, CB	6. Dir. 2D	7. Dir. 30 D
VVD	1.5%	2.4%	2.2%	7.9%	10.4%	11.8%	8.4%
PvdA	11%	17.5%	24.6%	48.7%	45.6%	.4%	7.9%
PVV	24.2%	4.5%	9.5%	6.4%	8.4%	0%	26.3%
CDA	4.4%	1.2%	1.3%	6%	4.7%	0%	4.2%
SP	1.8%	1.4%	1.3%	2.6%	3.8%	57.4%	21.5%
D66	9.6%	8.5%	8.4%	5.2%	4.8%	4.9%	5.6%
GL	9.1%	5.5%	5.6%	4.6%	6.09%	5%	13.8%
CU	31.3%	27.9%	26.3%	1.5%	1.4%	0%	2.5%
SGP	3.3%	.3%	.4%	1.8%	1.4%	20.3%	21.5%
PvdD	0%	14.5%	6.5%	6.3%	5.4%	0%	6.2%
TROTS	3.6%	16.2%	14%	8.9%	7.8%	0%	2.5%
Total	100%	100%	100%	100%	100%	100%	100%

1) Aggregate Consequences

I begin by looking at how much of an overall difference MAUD rule selection makes to the general output of our VAA. First I show the aggregate partisan recommendations generated by the different MAUD specifications (see table 6). In order to summarise the differences between these different MAUD rules for the aggregate outcome, I measure the (1979) Pedersen index of volatility generated by changing from one MAUD rule to another. The Pedersen index was originally intended to measure the extent of aggregate electoral volatility at the party system level between two elections, but here it is used to measure the extent to which aggregate a VAA's aggregate vote advice differs according to the specific MAUD in use, holding user and party issue positions constant. The Pedersen index ranges from 0% (identical outcomes) to 100% (where no parties who won votes in the first election win any votes in the second). In order to present these figures, I construct a pair-wise matrix of volatility between each of the 6 MAUD rules (Table 7). This analysis gives an estimation of how different the aggregate recommendations were from each other, viewed at the party system level.

When we first look at table 7, we see that it matters substantially which MAUD rule is used. There are particularly large differences in aggregate proportions of parties recommended between proximity and directional MAUD rules.

Table 7. Pedersen index of volatility for pairwise comparisons of 7 MAUD rules

	1	2	3	4	5	6	7
1	0%	34.4%	31.1%	58.1%	55.4%	82.9%	47.5%
2		0%	12.3%	46.1%	47.6%	85.4%	70.5%
3			0%	37.3%	36.8%	87.8%	65.25%
4				0%	7.3%	77.7%	59.4%
5					0%	74%	55.8%
6						0%	49.6%
7							0%

Particularly the contrast between proximity and directional theory approaches, especially when directional theory is applied in a two dimensional context. When one considers that a 100 per cent score indicates that two distributions are as different as can be, we find that index score for differences between proximity and directional (2D) MAUD's range between 74 and 87.8 per cent. Not so dramatic when directional accounts took in all 30 dimensions, however differences are still substantial. There are also strikingly different outcomes between low dimensional and high dimensional approaches, with volatility at around 50 percent between the two (comparing 1,2, and 3 with 4 and 5 in proximity and 6 with 7 in directional). If similar MAUDs and number of dimensions are used and the only variation is between Euclidean and city block metrics – the Pedersen index is very low (only 12.5 percent between 2 and 3, and only 7.3 per cent between 4 and 5).

2) *Partisan Consequences*

In this part of the analysis, I examine the partisan implications of MAUD rules for the parties competing in the 2010 Dutch elections. What most clearly emerges from summarising aggregate outputs in table 6 recommended is that directional models give disproportional large numbers of VAA users a recommendation to vote for an extreme party (SGP, PVV and SP). The proximity models give a disproportional large number of recommendations for the PvdA (centre left), the Christian Unions (centrist, religious party) and TROTS (an anti-immigration populist party).

In Table 8, I employ the stated 2006 vote choice of users as a baseline for comparison (accepting that many users may have changed their opinions since voting in 2006). As I argue in previous work (Wall et al., 2009), it can be counter-productive to compare VAA outputs to election results, as VAA users are not a representative sample of the overall population. Yet here the baseline merely serves to compare the seven MAUD rules.

When we first focus on the major Dutch parties, we see that under all MAUD rules the VVD is less frequently advised than their level of support among the site users in 2006. This is even more the case for the Christian democratic CDA, which has very low proportions of recommendations, compared to their level of popular support. While the centre-right parties are underrepresented in the number of recommendations given regardless of the MAUD rule, the social democratic PvdA has a very mixed record when it comes to the effect of the different models. Directional models give relatively few recommendations for the PvdA, while proximity models that impose 30 separate dimensions are very beneficial for the social democrats in terms of frequency of being recommended. The Socialist Party benefits from MAUD rules based on directional theory, due to their extreme position, while the proportion of recommendations under proximity model MAUD's for the SP are substantially lower than their actual vote share. The PVV benefits from all models, as they are more often recommended than their level of popular support.

Table 8. Proportion difference in first-preference vote advice, by party, compared to reported (2006) voting behaviour of site respondents.

Party	2006 party choice ³	1 Prox, 1D	2 Prox, 2D. Eucl	3 prox, 2D, CB	4 Prox, 30D Eucl	5 Prox, 30D, CB	6 Dir, 2D	7 Dir. 30D
VVD	17%	-15.5%	-14.6%	-14.8%	-9.1%	-6.6%	-5.2%	-8.6%
PvdA	20%	-9.0%	-2.5%	4.6%	28.7%	25.6%	-19.6%	-12.1%
PVV	4.7%	19.5%	-0.2%	4.8%	1.7%	3.7%	-4.7%	21.6%
CDA	17.2%	-12.8%	-16.0%	-15.9%	-11.2%	-12.5%	-17.2%	-13.0%
SP	14.5%	-12.7%	-13.1%	-13.2%	-11.9%	-10.7%	42.9%	7.0%
D66	8.9%	0.7%	-0.4%	-0.5%	-3.7%	-4.1%	-4.0%	-3.3%
GL	9.3%	-0.2%	-3.8%	-3.7%	-4.7%	-3.2%	-4.3%	4.5%
CU	5.6%	25.7%	22.3%	20.7%	-4.1%	-4.2%	-5.6%	-3.1%
SGP	0.7%	2.6%	-0.4%	-0.3%	1.1%	0.7%	19.6%	20.8%
PvdD	2.1%	-2.1%	12.4%	4.4%	4.2%	3.3%	-2.1%	4.1%
ToN	0.0%	3.6%	16.2%	14.0%	8.9%	7.8%	0.0%	2.5%

³ Counting only those users who reported voting for one of the 11 parties considered, n = 37,181.

These findings suggests that a parties' policy position, particularly their centrality or extremism causes these different effects of MAUD rules used for the aggregation of the recommendation. Also, MAUD rules that position parties in a two dimensional plane will discriminate parties that cluster together and more frequently recommended parties that have a spatially more unique position. Overall, however, it is clear that MAUD rules used to aggregate the issue stances of voters matter. Different MAUD rules generate dramatically different outcomes or recommendations which party is closest to these preferences. In the next section, I report qualitative research that focuses on the other key aspects of the VEA construction process (apart from MAUD rule choice), as exemplified by the 'KLieskompas method'. Thus I seek to give an overall picture of how VEA design and implementation influence to their contribution to the transparency and legitimacy of elections.

Analysis of overall methodology used in executing a VEA website

The purpose of this analysis is to describe the overall Kieskompas methodology, which involves eight key steps:

1. Selecting the most relevant political parties or candidates to be included.
2. Selecting authoritative sources – party platforms, policy documents, websites, debate transcripts and campaign material - from which party positions can be extracted, as well as creating a hierarchical order among different types of sources (should any two sources provide alternative positioning) for party placement.
3. Identifying the salient issues in a specific election year, i.e., selecting the issues on which political parties will be positioned, and on which voters will be compared to parties.
4. Framing propositions about these salient issues that give 'ideological' direction to the proposition and determine it's 'loading' on one of the deeper cleavage dimensions that demarcate the political landscape.

5. Searching the selected sources (as identified in step 2) and extracting specific text snippets that justify the calibrations of parties or candidates on an issue. Having selected these snippets, a party coding on each issue is decided by the coding team. Each specific coding decision at this stage is thus accompanied by the text snippets upon which that decision is based.
6. A process of party authorisation of coding decisions. This process begins with sending out the blank questionnaire to parties to collect their self-placements. The VEA team then compares their hand-coded placements on issues with the self-placements of the political parties. The team then undertakes a process of interaction with the parties until the maximum possible amount of agreement is reached over the coded position of the party (though the coding team retain discretion over the final coding decisions).
7. Comparing voters' opinions to issue positions of parties, i.e., summarising the voter-party comparison rules that drive the site's output. (this section is very short, as the issue is dealt with in such detail above in this report).
8. Developing a multidimensional 'map' of party competition (determining the salient cleavages underpinning party competition, and how the individual issue items load on these cleavages).

In the subsequent sections, I discuss each of these steps in turn, and outline some of the advantages and trade-offs inherent in the Kieskompas methodology.

Selection of relevant parties or candidates: Who's in and who's out?

VEAs can include all parties or candidates that participate in an election, however more typically they include only a selection of the most relevant parties. In multi-party systems with large numbers of parties registering for an election, filtering out micro parties that do not stand a chance of gaining parliamentary representation reduces the complexity of output for users. To avoid exclusion of electorally relevant parties, the Kieskompas method

uses two key criteria for inclusion: (1) all parties with one or more seats in parliament that enter the elections are automatically included; (2) parties that do not have any seats in the outgoing parliament but which consistently poll at least one seat in a number of opinion polls are also included. In candidate-based elections such as the US presidential elections (see www.electoralcompass.com), candidates were included when they consistently ended up in opinion polls as serious candidates with a genuine chance of winning the party's nomination or – adapting Sartori's (1976) notion of 'blackmail potential' – if they could take away the nomination from another candidate.

In the French presidential elections, for instance, candidates included polled at least one percent in opinion polls, while in Mexico all four candidates could be included. In countries with no electoral history or reliable opinion polling, such as in the founding elections of Egypt and Tunisia, where a wide variety of new parties exist, criteria for party selection must be adjusted. Common criteria in such contexts include: having formal recognition and legal status; fielding candidates in a sufficiently large number of constituencies to be nationally relevant; and being considered viable contenders by experts. In general terms, the trade-off in party selection is between the risk of over-inclusion (leading to an artificially crowded political landscape) and over-exclusion (leading to potentially relevant parties being excluded). As the latter is a more troubling outcome, the exclusion rules employed by Kieskompas only seek to exclude parties who are highly unlikely to receive meaningful parliamentary representation.

Selecting and analysing relevant party documents

Once the relevant parties or candidates are selected, VEA developers need to decide which sources provide authoritative information on issue positions and reveal the most salient issues on which parties compete. The most stable cross-time source for this is official documentation that parties or candidates themselves produce in order to communicate their policy positions (Budge, 2000; Laver and Garry, 2000; Marks et al., 2007; Mikhaylov,

Laver and Benoit, 2008). In most advanced democracies, such positions are formally documented in a manifesto that is especially written for the upcoming election. However, manifestos also have several drawbacks for VEA designers, and, indeed, for political scientists in general: they are often strategic (i.e., designed to please large groups of voters, rather than explain a party's true policy stances); they differ in terms of length and topics addressed across parties (important issues are not always addressed in all manifestos as parties tend to de-emphasise issues that can hurt them electorally); and the manifestos of government parties systematically differ from those of non-government parties, the latter being less constrained by recent 'track records' and possible future (coalition) government participation (Marks et al., 2007, pp.26-27). In addition, new parties may not even have a formal manifesto and even if they do have such a document, it may not reveal a policy position on all relevant issues. Indeed, some parties focus heavily on a single or a very limited number of issues. Even well-established parties which have agreed positions on a range of issues may only release their formal manifestos late in the electoral campaign, often too late to be useful for VEA designers – as VEAs are typically launched at the start of the formal campaign.

Thus, party manifestos are important, but they are not the sole source of party stances on issues. Party websites usually provide more extensive information on policies and additional information, including statements by leaders, press releases, motions officially accepted at party conferences and even full policy documents on specific topics, all of which helps VEA developers to pin down a party when an issue position is either ambiguous or not mentioned in the party manifesto. Additionally, the reactions of party leaders and representatives in the media can be a source for party positioning on issues. Party leaders' policy statements can be carefully planned speeches that are timed and targeted at specific voter groups or to get media attention. In addition, debates on radio and TV often serve to introduce new positions, to clarify stances taken or to sharpen the distinction between competitors.

If VEA developers use multiple sources, they need to determine which sources are relatively most authoritative and establish a hierarchy among them. This hierarchy may differ from country to country and even from party to party in the same country. Within traditional mass parties, the manifesto will outweigh any policy statement made by an individual party representative, whereas a declaration by the leader of a populist party may override the official party programme. A general rule for the relative weight of sources is that more recent sources should be given precedence; so current manifestos outweigh previous manifestos, recent speeches by the leader outweigh those made in the past, and so on. This hierarchy of documents should be formalised (*i.e.*, a rank ordering of sources should be written up for the coding team) and integrated into the coding template. With certain sources, some prudence is required. In parliamentary debates, for example, representatives of government parties may not express the official position of the party, but the government's position on the issues, or at least they may formulate their contribution to the debate in a manner that is acceptable to coalition partners. Even more caution is required with the statements of party leaders when such leaders are also government ministers. VEA developers should be very careful in distinguishing in what capacity and setting political statements are made. Even during debates, parties that are often in government will make declarations with current and future coalition bargaining in the back of their heads. As a consequence of their coalition potential, such 'governmental' parties are much more difficult to position on issues than minor, more extreme and permanent opposition parties. This is why it is advisable to enter into an authorisation process with the parties themselves, which will be discussed below.

Issue identification and selection: Determining what's at stake in the election

The selection of the issues to be included in the VEA is the most crucial choice to be made in the construction of the application. Walgrave et al. (2009) ran a large number of simulations in which a subset of issue statements was taken from a larger pool of items on which the positions of parties and users were known. They found that the statistical dispersion of the matches observed between party and voter is relatively narrow, but they also underline that

in real political terms these differences are really substantial, concluding that issue-selection dramatically affects site output. At the moment, a major problem for VEA designers is that there is no current consensus among political scientists as to how salient issues can be identified. In order to limit political mismatches, issue identification should be the result of a process involving several methods and actors.

Although extant methods for statement development differ substantially, most party profilers start by drawing up a list of relevant issues, and subsequently narrow down this list to a final selection (Laros, 2007; Deschouwer and Nuytemans, 2005; van Praag, 2007; Krouwel and Fiers, 2007, pp.7-10). Statements can be extracted from the party programmes or websites, using automated text analysis or careful reading, but can also be drawn up by country experts and/or journalists. The Kieskompas method combines a computer assisted text analysis with expert opinion. Kieskompas collects the relevant party platforms and website content and then uses large corpus text analysis methods to extract salient issues in terms of frequency. In the most recent Kieskompas iteration for the 2012 Dutch legislative elections, issue extraction from manifestoes was complemented by an analysis of relative lemma frequencies within the full content of main newspapers over the last legislative period (2010-2012). For this extraction, the entire text corpus of media-content was lemmatized and tagged and the relative frequency of each lemma was compared to a reference corpus, using Latent Dirichlet Allocation (LDA). LDA can be seen as a latent variable model, with the latent parameters being the topics per document (Blei, NG, and Jordan 2003; Griffith and Steyvers, 2004). This method extracts latent factors (clusters of topics) that indicate which issues were dominant in the public discourse and how these issues were discussed.

These computer-generated issue-lists are complemented with two further issue lists, one produced independently by the country experts and one by journalists. Kieskompas asks academic country experts to carefully read the party programmes and other relevant text sources to identify salient issues, not simply by their frequency, but on a scale of 'urgency' or 'strength' of the issue. This approach prevents the exclusion of issues that are mentioned

less frequently but are phrased in very strong terms, and can therefore be considered important for a party. Journalists compile a list of the most relevant issues based on their day-to-day observation of national politics. Both experts and journalists are asked to indicate the core of the political dispute at stake, information that is used in framing the proposition.

All issue lists are then compared and the most frequently mentioned and the most strongly emphasized issues are included in a list of roughly 35 to 55 issues, drawn up in multiple sessions with the country experts. Although Kieskompas normally uses 30 issues, parties and candidates are positioned on a wider range of by the coding team in order to allow experts exclude issues on which party positions cluster too much to usefully discriminate positional differences. The selected 30 issues need to be salient for political competition, address a wide range of relevant policy fields in order to prevent a bias towards specific parties that 'own' certain issues and, to make the final cut, they must be able to differentiate between parties. Issues on which parties broadly have shared goals should be avoided. These 'valence issues' or 'empty signifiers' touch on such broad and primary issues where parties only differ on the ways to achieve the same goal (Walgrave, 2008). The involvement of numerous experts for selecting salient issues combined with computer-assisted content analysis is designed to minimise bias in the issue selection and is a substantive empirical improvement over *a priori* selection of issues and bipolar policy dimensions. However, the absence of a universally agreed process for statement selection remains troubling for VEA practitioners, particularly in the light of Walgrave et al.'s (2008) research findings.

Framing and 'scaling' the issue propositions

One of the most contested issues in VEA design is how to frame the issue propositions in such a manner that they can be used to elicit voter and party positions without introducing bias. VEAs differ considerably in their approaches to the development of specific propositions. Some VEAs take the statements directly from the party platforms or invite

party representatives to co-develop the statements during workshops, while others only involve experts in the formulation process. The framing of these statements also depends on the answer categories and whether their propositions need to be scaled to deeper-lying issue dimensions.

While many VEAs use binary answer categories (such as 'yes/no'; or 'agree/disagree'), Kieskompas uses Likert items with five response levels. The answer categories can vary, so that also propositions can be framed towards a five-point scale ranging from strongly agree to strongly disagree; from much more to much less; from much lower to much higher, etc. The Likert item format was adopted after studies found that Dutch VEA *Stemwijzer's* binary answer categories ('agree' versus 'disagree', as well as a 'don't know') denied parties and voters the opportunity to adopt nuanced positions and resulted in incorrect (self-) positioning of parties on multiple propositions, which distorted the voting advices given (Groot, 2003a, pp.23-24). Dichotomous answer categories also reduce the discriminatory power between parties belonging to a similar political orientation as well as of centre parties (Groot, 2003a, pp.24-27; Arendsen, 2003; Kleinnijenhuis and Krouwel, 2007, 2008). With Likert items, parties and users can indicate the direction of their attitude, as well as the intensity to which they (dis)agree with the proposition. This allows for a more nuanced differentiation between parties, even when they belong to the same political 'camp'. The labelled five point scales also offer users and parties a centrist position. For political parties this is important, as the centre is a real existing political position. When VEAs like *Stemwijzer* and *Smartvote* do not use the middle category, they force parties and voters to take sides, even when they are actually centrist or neutral on the issue. Nonetheless, the decision to use Likert response scales, and especially the provision of a 'neutral' midpoint, implies some tradeoffs, especially for particularly controversial political issues where a 'neutral' response may be a safe haven for satisficing voters, or for parties who do not wish to engage with an issue (Gemenis, 2011). How the middle category is interpreted by users, which statements drive respondents to the middle category, whether the midpoint indicates an non-attitude or a centrist attitude and what the best label is for this mid-point is still fiercely debated (see Triga et. al 2012).

One obvious advantage of using Likert items over binary answer categories is that issue propositions can be considered items of more profound political dimensions, such as the left-right or conservative-progressive divide. Locating parties or candidates in a common political space brings us to the problem that there is no consensus over the number and type of salient cleavage dimensions, let alone on which issues belong to which deeper-lying ideological dimension. When issues are scaled towards cleavage dimensions, two problems need to be addressed. First, determining to which dimension an issue needs to be scaled and secondly the weight of each issue on this dimension. To avoid bias towards one ideological orientation, party or candidate, propositions in a VEA need to be framed in different 'political directions'. The entire set of propositional statements must be 'balanced' in terms of subjective framing. Some VEAs balance the propositions by taking an equal number from all party platforms. Within the Kieskompas method, roughly a quarter of the questions will have 'rightist' framings, a quarter 'leftist', a quarter 'progressive' and a quarter are formulated in the 'conservative' direction. As such, each proposition is scaled towards one of the four poles of the political landscape. The total set of propositions is also grouped into broader issue clusters or topics, such as healthcare, environmental policy, welfare state, foreign policy, etc. Within such broader topics, several propositions are thus 'framed' in a variety of 'ideological directions' in order to reduce the danger of acquiescence bias, 'pushing' the voter towards a certain response (Evans et al., 1996). Also, clustering multiple propositions reduces the effect of the selection of specific issues and allows for more subtle levels of differentiation between parties, particularly when they are in the same political block or in the crowded political centre.

Both the selection and directional framing of issue propositions are vital to the functionality and fairness of the tool – however, we must acknowledge that issue selection and proposition framing are human work. Even if it is done in the most rigorous manner possible, it remains a qualitative and interpretive process. However, bias can be minimised by using a range of diverse sources and multiple experts who balance the issue selection and framing to prevent design pitfalls of proposition formulation (van den Berg et al., 2002). A further difficulty of dimensional framing is that occasionally questions are framed in a double-barrelled manner, as observed by Gemenis (2011). However, in many instances,

such phrasing is necessary to convey the interdependency of dimensional issues – for instance, the necessity to pay for promised increases in public spending with increases in government tax take. For example, while the statement ‘Social programmes should be maintained even at the cost of higher taxes’ from the 2009 EU Profiler site may be double barrelled, this type of formulation is necessary to convey the trade-off between taxation and social spending.

With regard to weighting, VEA-developers could opt to simply weigh all issues as 1, or assign a relative weight to each issue. For this relative weighting, as well as determining if the issue is scaled to the correct dimension, data can be used from previous iterations or from elections studies and surveys that included questions on the concerned issues. While VEA designers have typically used *a priori* dimensional assignment (according to the judgement of country experts), both the dimension to which an issue scales and the weight of its loading to the deeper-lying dimension can be determined through factor analysis (Krouwel and Kleinnijenhuis, 2008). VEA developers could also opt to scale an issue to two dimensions. Experiments are currently being carried out (van der Linden et al., 2012) employing inductive dimensional analysis of public opinion data to construct weightings that are based on the extent to which individual items scale on a given factor. This represents a promising approach as Gemenis (2011) provides evidence that inductive dimensional analysis often points out that often items that are scaled to a dimension *a priori* by coding teams do not scale well empirically. An alternative approach allows voters to exclude issue categories from their output calculation (i.e., giving items that fall into such issue areas a zero weighting).

Placing the parties

Political scientists devote considerable time, effort and expense to the study of parties’ policy stances and they have developed several methods to determine the ideological or policy positions adopted by parties. Examples of such methods include: expert surveys;

surveys of placement of parties by voters; surveys of the political parties themselves; roll-call behaviour of politicians; and content analysis of party manifestos (for a detailed discussion, see Benoit and Laver, 2006, pp.123-153). VEA developers use various methods to calibrate parties on the issues (Gemenis, 2011). One method (used for example by Dutch VEA Stemwijzer) is simply to allow parties to position themselves on the issues. A second method is to let experts position parties or candidates on the issues, either with or without analysing party documents. A third method is automated coding of the political texts by computer programs to determine party positions (Benoit and Laver, 2003, 2007). There are, however, major methodological problems with automated text-analysis. When 'word-based' techniques are used and the distribution and patterns of words in text are analysed, this does not tap into the connotation of words in their context – crucial in political communication – and they chunk texts into words, negating semantic, grammatical and discursive structures and frames that give meaning. Moreover, automatic text analyses are incapable distinguishing positive or negative direction in a text, which is crucial for identifying relative party positions on issues. Considering the importance of VEAs we have to accept that hand coding is an interpretive approach, which is very labour-intensive, expensive to replicate or change and sensitive to individual errors of interpretation and low inter-coder reliability (see Budge et al., 1987; Klingemann et al., 1994).

Letting parties position themselves is a method that is susceptible to manipulation by political parties (van Praag, 2007), while experts can also be biased if no control mechanism is included. In 2006, Kieskompas developed a method that combines expert judgements based on text-analysis with self-placement of parties, including an interactive process with the parties. An important improvement on previous experts' surveys and safeguard against bias is that experts base their judgements on thorough text analysis (instead of general interpretations) and fully document and justify each positioning on an issue with text snippets from official party documentation. Thus, unlike most expert surveys and VEAs – Vote Compass is completely transparent with regard to source texts and coding. In order to maximise transparency, all text extracts underpinning the party calibrations are made available to each user with a simple mouse-click, allowing voters to also enter into a debate on the position of the parties on these issues. Improving further on expert surveys, the

Kieskompas method fully documents and justifies each positioning of parties or candidates with text snippets from official party documentation. Thus, unlike traditional expert-surveys, the Kieskompas method leads to a database of party positions on salient issues in which each position attributed to the party is justified by a reference and hyperlink to a text snippet from an authoritative party document. All text snippets that are used as evidence for the calibration of parties' positions are added to the Vote Compass website and made accessible with hyperlinks, so that the positioning of parties and candidates is completely transparent, also to the user. In addition, Vote Compasses provide hyperlinks directly to the *original* document that was used to position the party. The manner in which this information is displayed in Figure 2, and it is important to note that each text snippet is accompanied by a hyperlink that allows the user to examine the original source document (and thus to check whether the snippets may have been taken out of context).

Figure 2 – Output displaying text justifying a single issue coding (note hyperlink to source document at the bottom of this text). Example taken from Mexican 2012 Vote Compass: <http://www.bruijulapresidencial.mx>

The screenshot shows the 'Brújula Presidencial' interface. At the top, there are navigation links: Main, Questions, Candidates, and Result. Below this is a 'close' button. The main content area features a profile picture and name for Gabriel Quadri de la Torre. The issue being discussed is 'Economy and Nationalism' with the specific statement: 'The government should be the only one responsible for producing and refining oil'. There are two rows of response buttons: 'Your answer' and 'Opinion'. Both rows have six options: 'Completely agree', 'Tend to agree', 'Neutral', 'Tend to disagree', 'Completely disagree', and 'No opinion'. In the 'Your answer' row, 'Neutral' is selected. In the 'Opinion' row, 'Tend to disagree' is selected. Below the buttons is a 'Justification' section containing a text snippet and a hyperlink: 'http://www.nuevaalianza.mx/plataforma.php#apertura'.

Notwithstanding elaborate safeguards in the various methods applied, pure objective party calibration will remain impossible as language production and interpretation is necessarily inter-subjective. The calibration of parties and candidates is also complicated by other problems. Sometimes no position can be found in official texts for some parties. If

then the party does not want to identify its stance on the issues by self-placement (see next section), the party is coded as 'no opinion', for want of a better solution. Another problem encountered by VEA developers is that parties are not always unitary actors. When parties are internally divided into factions, it may happen that individual members or different documents of the same party state dissimilar policy positions. In such cases, Vote Compass usually opts to show all the different positions, including the text-snippets, in order of the 'official' or 'dominant' position first, followed by the deviant positions. Thus, there may be controversy over which and how many text sections should be included, as well as where to cut off the text snippet. In short, there are many reasons why positioning parties or candidates may be less than straightforward. In the next section we describe the process of authorisation, which involves contacting the parties or candidates and locking them into a clarification of their stances.

Contacting the parties for self-placement: Reaching agreement

Kieskompas brings an important innovation to party profiling methods in that it adds a self-placement by each party or candidate to the text-based coding by experts. Country experts contact actors within each party who are authorised to assess that party's stance on a given issue. For some parties, the party leader will fill out the self-placement questionnaire, while in others authorised party spokespeople or campaign leaders will fill out the form. In some countries, parties now even assign a team with the task of reacting to the various VEA-requests. Self-assessment of their issue position gives parties the opportunity to challenge expert-coding decisions. Each party is provided with the blank propositions at an early stage and are also asked to position themselves and provide appropriate text snippets from their manifestos or other official documents to substantiate and justify their self-placement. Self-placements are then compared with the hand-coded positioning data and well-considered adjustments can be made. Discrepancies between the self-placement and coder positioning are communicated to the parties, as well as the text-snippets that the coders used as a justification. Parties are subsequently asked – in extensive direct communication - to

indicate what they consider to be incorrect about the disputed expert judgement and are allowed to provide alternative text-snippets in order to substantiate their self-placement.

This interactive process with parties is generally (with a few important exceptions) constructive and leads to clarification where issue positions were unclear. For example, in the Dutch elections of 2006 the hand-coded calibrations of Kieskompas corresponded with 82 per cent of the party self placements, while after two rounds of contact this level of agreement increased to 92 and 99 per cent respectively (Krouwel and Fiers, 2006, p.22). Indeed, most of the disagreement is related to the intensity of the calibration, not of the direction. Usually, after a first round of discussion the agreement-level increases to over 90 per cent, while multiple rounds usually result in only very few propositions remaining disputed between the expert and the party itself. Willingness to participate in the self-placement procedure varies across countries (Wagner and Ruusuvirta, 2011, pp.5-6), and depends on the authority of the expert(s) contacting the party and the number of iterations and success of previous VEAs. Kieskompas usually achieves compliance from a majority of parties, while a cross-national project during the EU elections in 2009 showed lower levels of participation (Krouwel and Fiers, 2007; Trechsel and Mair, 2011). A major difficulty with asking parties to self-place is that parties face strategic incentives to misrepresent their issue positions in order to 'game' popular VEA sites (Ramonaite 2010; Wagner and Ruusuvirta, 2011).

However, the methodology described above seeks to control against much manipulation firstly by requiring parties to provide formal texts justifying their self-placement, and secondly by leaving the final coding decision in the hands of the VEA team for unresolved coding disputes with parties. Gemenis (2011) points out that expert coding is often low in terms of inter-coder reliability. His proposed approach of checking inter-coder reliability to avoid inter-coder dynamics leading to distorted codings is a valid one, and should be considered for future VEA iterations (resources permitting, of course). However, the process of engagement with parties is highly beneficial as they are the actors who find their political positions being assessed by VEA designers. The process is designed to deal with arbitrary

expert judgments and allows parties to respond to potential miscoding without handing them complete discretion over their placement.

Comparing Voter and party positions: what approach to take?

The issue of MAUD rule selection and its consequences for VEA outputs are discussed in detail in this report – I note here only that such selection is a core part of VEA development, and, as we have seen, it is a substantively important choice. The Kieskompas method relies on a two-dimensional spatial conception of political competition – which results in a spatial ‘map’ of a political system, upon which the voter is located. I discuss the process by which ‘maps’ are generated in the next subsection.

Developing a multidimensional ‘map’ of party competition

Using confirmatory factor analysis and Mokken scaling analysis, Kieskompas check whether each of the key election issues identified can be scaled to the overarching dimensions. If an issue does not ‘scale’ on either dimension, it can be excluded from the algorithm used to generate the political ‘map’ and users can be encouraged to consider this issue in isolation. The formal Vote Compass calibration procedure computes the summated positions $\sum_k P_{ojk}$ of each party o and averaged positions of the voter $\sum_k I_{jk}$ on the two dimensions j_1 and j_2 , before distances along each of the axes are computed. The averaged positions of the parties and the user are presented graphically along the two dimensions, thus giving rise to Euclidean distances between a voter and a party.

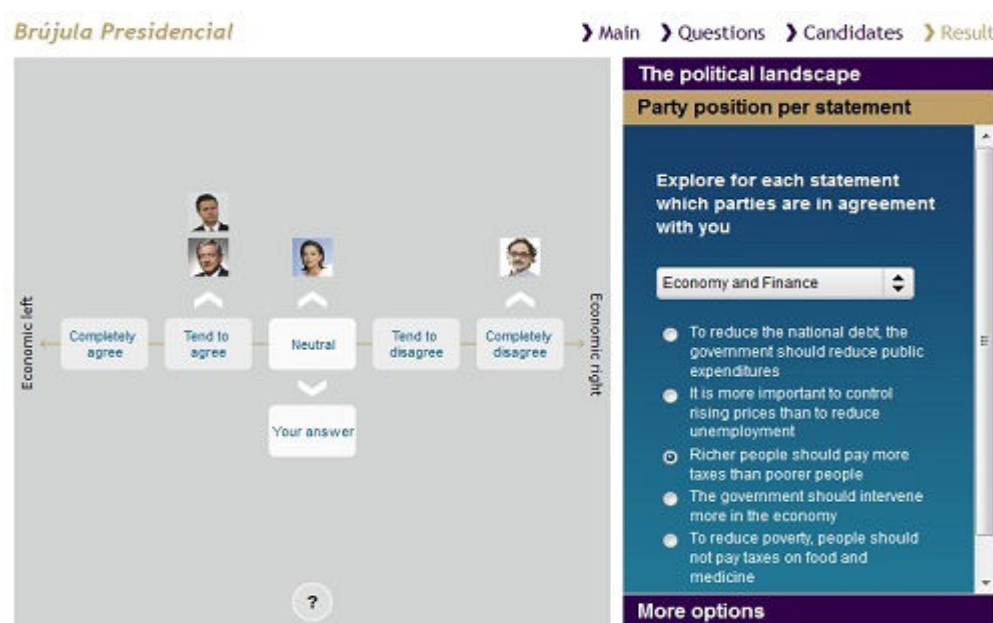
$$A_o = \left[\sum_j \left| \sum_k P_{ojk} - \sum_k I_{jk} \right|^2 \right]^{1/2}$$

The site also visualises the extent to which users’ spatial position is made uncertain due to ‘ideological inconsistency’ (i.e., the extent to which their answers deviate from the logical

pattern anticipated by our scaling procedure) via an ellipse which is drawn around the user's issue position – based on the standard deviation of their answers on the x and y axes. This is a deliberate addition, and a critical component of the placement of VEA users in a political space. The 'position' of the user is not one single point, but an approximate area which will be wide for ideologically inconsistent voters, whose answers are a mixture of left and right, progressive and conservative, and narrow for voters whose answers point consistently in the same ideological direction.

The visual representation of an approximate area also allows the users to observe their relative position within the overall dimensions represented. An example of this output is provided in Figure 1 above. It is also important to point out that well-designed VEAs facilitate multiple outputs – and their central idea is to encourage voters to examine party policy positions. As well as the political 'map' and 'closeness' outputs, Kieskompas sites also allow users to compare their responses with party positions on each individual issue measured by the application. Thus, while producing a 2 dimensional summary of the political space, the Kieskompas design also seeks to cater for users who are interested in a given single issue or in only a subset of issues. Figure 3 displays this issue-by-issue output.

Figure 3 – An example of single issue output on the question 'Richer people should pay more taxes than poorer people' (translation); with voter 'no opinion'. Example taken from Mexican 2012 Vote Compass: <http://www.brujulapresidencial.mx>



This analysis has sought to show how the execution of a VEA site is a multistage process, with site designers making important theoretical and methodological choices at each stage. In terms of addressing key question 2, this analysis has sought to show firstly that variations in the type of MAUD rule impact strongly on VEA outputs, with observable partisan implications. This is a troubling finding in terms of the contribution of VEAs to the legitimacy and transparency of elections, as the ‘advice’ that they provide to voters is heavily influenced by a choice among competing MAUD rules – with no clear guidance from the literature as to which rule is ‘best’. The analysis also provide a clear account of the ‘Kieskompas method’ with a view to facilitating the analysis of how Kieskompas websites, a phenomenon of the modern global communications environment, can best contribute to the legitimacy and transparency of the electoral process.

- (Key question 3) What can data gleaned from Kieskompas email surveys tell us about online survey response rates?

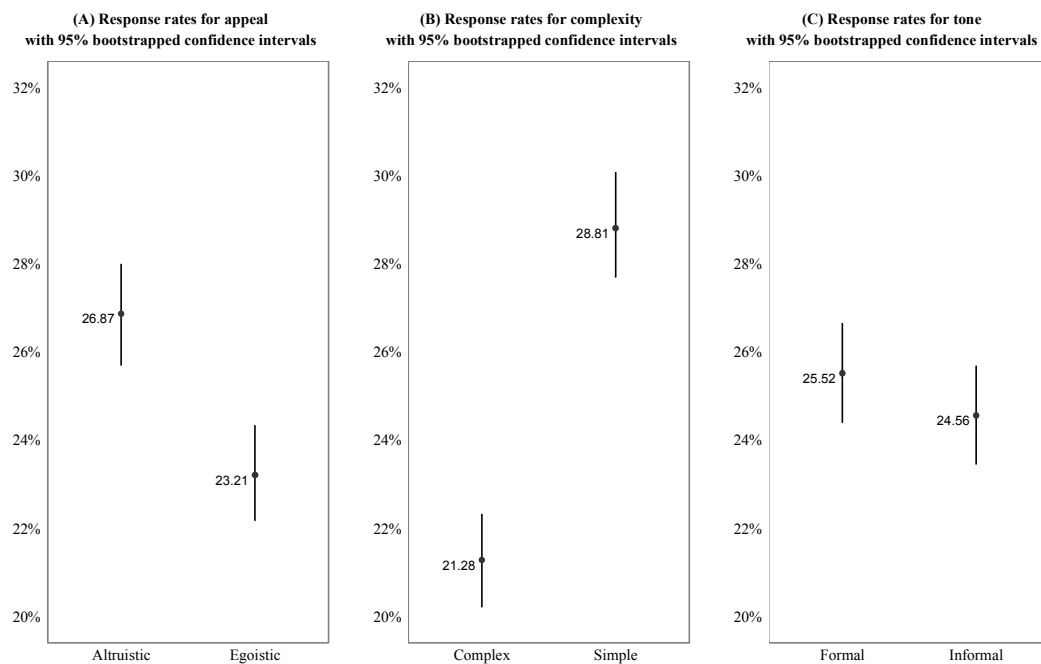
Response Rates

Following a classic approach when dealing with different treatment groups, we are interested in the average treatment effect associated with our manipulation – the choice of the combination between tone, content, and difficulty. The limitation of this approach is that the initial sample of around 11, 300 respondents already makes use only of those people, who once supplied their email address and voiced their interest in participating in upcoming surveys. Thus, our findings should not be extended for low participation propensity individuals, as they already dropped out from the sample.

Figure 4 displays the average response rates for each experimental group (with 95% confidence intervals). On average, altruistic appeals generated higher response rates and these differences between egoistic and altruistic appeals are statistically significant altruistic appeals register higher response rates on average, compared to egoistic messages. On both sides of the appeal dimension, complex messages perform badly. However, if they are

paired with an altruistic appeal the situation is not as grim, but the response rates are still low. The egoistic-complex-informal message has the lowest response rate, 15%, which is also statistically significantly lower than the rates for any other group.

Figure 4. Response rates for each group with 95% bootstrapped confidence intervals.



Analyzing the treatment groups we see no significant difference in how these treatments affect the composition of respondents on the previous contacts covariate. These results suggest that for the present survey the decrease in response rates stems from the loss of individuals who already participated in three surveys administered by Kieskompas, suggesting a sort of survey participation *apathy* or *saturation*.

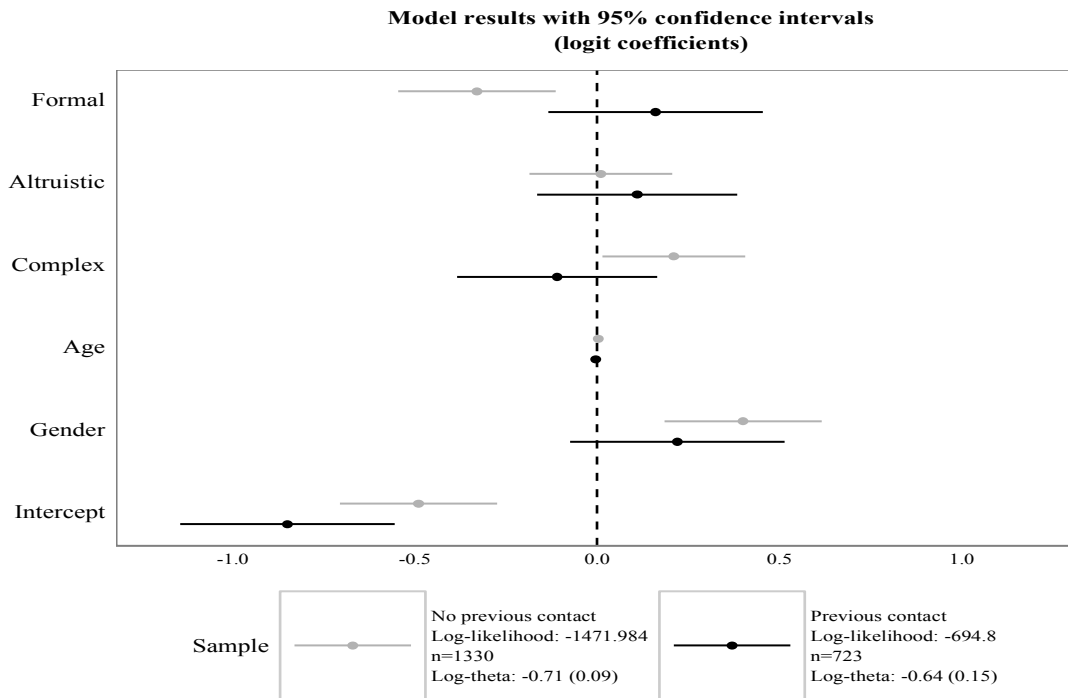
From a practical perspective, a tentative conclusion from this step would be to keep the message simple and appeal to altruistic motivation. These findings are in line with results reported by Gendall et al (1995) on mail cover letters. Also, if this is not possible or the messages are medium difficulty, altruistic appeals work much better in attracting respondents. Furthermore, we saw that all these messages fair best in attracting relative

newcomers, and they all do a relatively poor job in convincing those individuals who previously were extensively re-contacted. We turn to response quality issues in the next subsection.

Response Quality

I now concentrate on the quality of responses that we received. In this case, we were interested in analyzing between-individual differences, examining the treatment effect associated with the re-contact email messages. As a first indicator of response quality I use item non-response. We tagged the “don’t know” answers as item non-response for the questions included in this survey. Not just it is a generally accepted to treat these non-responses as missing values in the analysis, and thus making it important to understand what determines them, the implementation process also gives us substantial reasons. The questionnaire was completed on a website, where in order to proceed the question had to be answered. Accordingly, any non-response is recorded as “don’t know”. For this indicator, first we can simply compare the average level of item non-response among groups as a first step. Figure 5 displays the group average item non-response (as count for all the items) with 95% confidence intervals. As expected, since all of the respondents recruited for this experiment were opt-in volunteers, the non-response counts are very low. Furthermore, there are only small between-group differences and these are not statistically significant.

Figure 5. Item-missing count with 95% confidence intervals for all groups.



5. Summary & conclusions

In recent years, millions of voters across many countries consulted a VEA website (or, in some cases, several such sites) prior to casting their vote, especially in national legislative elections in European multiparty systems. VEA websites were pioneered in the Netherlands, where an early text version appeared just before the dawn of the internet age, in 1989, followed by the launch of the *Stemwijzer* (which roughly translates as ‘Vote Pointer’) website in 1998. However, VAA sites have since become internationally widespread: an expert survey of European political scientists revealed that VAA sites featured in national elections in 18 European countries. In countries where VAA sites have now featured in several elections and consistently generated high volumes of unique site visits, Walgrave *et al.* (2008) conclude that ‘voting advice has become a natural part of the campaign’ (p. 43). Nowhere is this more true than in the Netherlands. While the 1998 Dutch *Stemwijzer* site generated approximately 250,000 recommendations, the 2006 version generated 4.6 million, with a further 1.5 million recommendations generated by the *Kieskompas* (Electoral Compass) site. The 2006 Dutch Parliamentary Election Study (DPES) reveals that these figures do appear to have corresponded to massive usage of such sites by the Dutch population – they report that 38 per cent of voters in the DPES survey sample recalled visiting a VAA site during the campaign.

More recently, VEA sites have been developed for the non-European elections, including the 2008 US presidential elections (<http://www.electoralcompass.com>), the 2011 Canadian Legislative elections (www.votecompass.ca), and series of websites in ‘Arab Spring’ elections in Tunisia, Morocco and Egypt (see: <http://www.kieskompas.nl/>). VAAs have also been deployed during both sub and supra national campaigns – the 2009 European Parliament elections featured a VAA that was consulted by over a million users (euprofiler.eu), while local and waterboard elections have been developed in the Netherlands, where the concept originated in the late 1980s. Thus as VAA sites are an increasingly widespread feature of electoral politics in the Internet age.

In this report, I have grouped my work on VEA sites during my time as an ELECDEM ER at Kieskompas under three key questions, each of which relate to the overarching themes of the ELECDEM project. In this conclusion, I will briefly recap my main findings, before reflecting on the grounds for and potential direction of new research in this exciting nexus of research and practice. Firstly, I examined whether VEAs could be said to have a political impact of their users, focusing on whether the ‘recommendation’ received by users exercises a discernible influence on their vote choice. Using an innovative research design that leveraged the temporal ordering of a series of survey questions, I demonstrated that, for the sample that I analysed, a clear ‘VEA effect’ is discernible; VEA users are more likely to vote for parties that are recommended to them online, although this effect is only visible when the VEA recommends a party that the user was seriously contemplating in the first place.

It is worth inserting some caveats about my findings under key question 1 at this point. The survey analysed does not benefit from the statistical and analytical advantages of being based on a randomised selection of site users. In order to maximize the number of participants, the option of signing up for future surveys and leaving an email address was made open to all users. In the event, 8,125 site users left viable emails and gave permission to be contacted post election, and just over 50% of these users completed the questionnaire. The site log file recorded over 800,000 unique visits during the campaign, so this self-selecting sample represents a tiny portion of the population of site users, who were not randomly selected. These limitations of our data make it impossible to extrapolate from my findings to the Dutch electorate as a whole with confidence, and we will therefore not attempt to do so here. A major improvement to be sought in future research on this subject is thus to seek to get an overall estimate of VEA impact. It is also worth considering that VEA impact may vary across countries and elections – a multi-country, multi-election could therefore provide considerable insight for future studies.

Furthermore, in concluding that VEAs have an impact on their users, I have to take account of the adaptive capacities of VAA designers, who may choose in future to downplay the

'advice' elements of their sites, which are typically designed with citizen education, rather than citizen persuasion as their goal. As I explained in analysing key question 2 – VAA designers make a series of choices in implementing their sites and changes in these choices away from an 'advice' site and towards an 'engagement' site may mean that future VEAs/VAAAs exert less of an active influence on vote choice than the site studied here. Indeed, the unclear nomenclature for describing VAAAs/VEAs reflects the current uncertainty about the precise function and optimum design of such sites. Certainly, the analysis of the impact of MAUD rule selection provided in my investigation of key question 2 indicates that VEA designers should be careful not to overstate the certainty of their results to their users, if they wish to maximise their contribution to the transparency and legitimacy of the electoral process. Again, future research on this topic could expand across datasets from multiple countries to develop more solidly generalised intuitions about how certain types of policy stances are favoured or punished by different MAUD rules.

The experimental data gathered to assess key question 3 provided some interesting insights into the importance of email wording for online survey participation requests. We found that requests written in simple language and appealing to respondents' altruistic motivations were more effective in eliciting responses than requests using complex language or relying on egoistic appeals. However, this study represents a tiny proportion of what is possible with the sort of data collected by VEA sites – especially when these data can be supplemented with follow up surveys at a relatively low cost. The super-large size of the data mean that they can be used to examine phenomena at a far more fine-grained level than typical surveys, which have a far smaller n. Furthermore, VEA data are time stamped throughout the campaign, meaning that they represent a hugely promising source for the examination of campaign dynamics. In my opinion, the future of VEA research will involve a number of innovations in the treatment and use of VEA generated data that allow it to contribute to the analysis of core questions in political science, and to take a place among the methodological mainstays of election studies.

References

- Aarts, K. and van der Kolk, H. (2007). 'The parliamentary election in the Netherlands, 22 November 2006', *Electoral Studies*, Vol. 26, pp.797 - 837.
- Andeweg, R.B., Irwin, G. (2002) *Governance and Politics of the Netherlands - Second Edition*. Houndmills: Palgrave MacMillan.
- Arendsen, R. (2003). 'Partijen zonder mening winnen op Stemwijzer.nl'. In: *Politiekdigitaal* (online). Available at: http://www.politiekdigitaal.nl/nieuwedemocratie/partijen_zonder_mening_winnen_op_stemwijzer.nl
- Arian, A., Krouwel, A.P.M., Pol, M. & Ventura, R. (2010). 'The Election Compass: Party profiling and voter attitudes'. In A. Arian & M. Shamir (Eds.), *The Elections in Israel 2009 (The Elections in Israel since 1969)*. New Jersey, USA: Transaction Books, pp. 275-298.
- Benoit, K. and Laver, M. (2003) 'Estimating Irish Party Positions Using Computer Wordscoring: The 2002 Elections', *Irish Political Studies*, Vol. 18 No.1, pp.97 - 107.
- Benoit, K. and Laver, M. (2007) 'Estimating party policy positions: Comparing expert surveys and hand coded content analysis', *Electoral Studies*, Vol. 26, pp.90 - 107.
- Benoit, K. and Laver, M. (forthcoming). "The Dimensionality Of Political Space: Epistemological And Methodological Considerations." *European Union Politics*.
- Benoit, K. and Laver, M. (2006) *Party Policy in Modern Democracies*, London: Routledge.
- Benoit, K., Laver, M. and Mikhaylov, S. (2009). 'Treating Words as Data with Error: Uncertainty in Text Statements of Policy Positions', *American Journal of Political Science*, Vol. 53 No.2, pp.495 - 513.
- Berg, H. van den, Mehcziz, M., Houtkoop-Steenstra, H. and Holleman, B. (2002) *Opinie- maken of opinie-meten. De rol van stellingvragen in markt- en opinieonderzoek*, Stichting voor Culturele Studies, Amsterdam.
- Biemer, Paul. (2010) "Total Survey Error: Design, Implementation, and Evaluation." *Public Opinion Quarterly* 74(5): 817-48.

Blei, David, Ng, Andrew & Jordan, Michael (2003), "Latent Dirichlet allocation". *Journal of Machine Learning Research* 3 (4–5): 993–1022.

Brennan, Mike. 1992. "The Effect of a Monetary Incentive on Mail Survey Response Rates: New data." *Journal of the Market Research Society* 34 (2): 173-77.

Breuer, F. (2010). 'The EU Profiler: A new way for voters to meet parties and to understand European elections', in Gagatsek, W. (ed.), *The 2009 Elections to the European Parliament - Country Reports*, European Union Democracy Observatory, Florence, pp. 27-31.

Budge, I. (2000) 'Validating Party Policy Positions', *British Journal of Political Science*, Vol. 31 No. 1, pp. 210-223.

Budge, I., Robertson, D. and Hearl, D. (eds.) (1987) *Ideology, Strategy and Party Change: Spatial analysis of post-war election programmes in 19 democracies*, Cambridge University Press, Cambridge.

Çarkoğlu, A. and Hinich, M. J. (2006) 'A spatial analysis of Turkish party preferences'" *Electoral Studies*, Vol. 25 No.2, pp.369 - 392.

Çarkoğlu, A. and Krouwel, A. (2012) *Political Space and Ideological Constraints: Evidence from Voting Advice Application Data in the Turkish Elections of 2011*, paper presented at the Workshop on Voting and Party System in Turkey's June 2011 Elections, Yasar University Izmir 9-10 February.

Carlson, T., Strandberg, K. (2005) 'The 2004 European parliament election on the web: Finnish actor strategies and voter responses', *Information Polity*, Vol. 10 No. 3, pp.189 - 204.

Cedroni, L. (2010) 'Voting Advice Applications in Europe: A Comparison', in Lorella Cedroni and Diego Garzia (eds.) *Voting Advice Applications in Europe: The State of the Art*, ScriptaWeb, Napoli, pp.247 - 257.

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

Chang, Linchiat, and Jon A. Krosnick. 2009. "National Surveys via RDD Telephone Interviewing vs. the Internet: Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly*, 73: 641-78.

Christian, Leah Melani, Don A. Dillman, and Jolene D. Smyth. 2008. "The Effects of Mode and Format on Answers to Scalar Questions in Telephone and Web Surveys." In Lepkowski,

- J., Tucker, C., Brick, M., de Leeuw, E., Japac, L., Lavrakas, P., Link, M., Sangster R. (eds), *Advances in Telephone Survey Methodology*. New York: Wiley, 205-75.
- Costa Lobo, M., Vink, M and Lisi, M. (2010) 'Mapping the Political Landscape: A Vote Advice Application in Portugal', in Lorella Cedroni and Diego Garzia (eds.), *Voting Advice Applications in Europe. The State of the Art*, ScriptaWeb, Napoli, pp.143 - 185.
- Couper, Mick P. (2000) "Web Surveys: A Review of Issues and Approaches." *Public Opinion Quarterly* 64: 464–94.
- Couper, Mick P. (2011) "The Future of Modes of Data Collection." *Public Opinion Quarterly* 75(5): 889-908.
- Curtin, Richard, Stanley Presser, and Eleanor Singer (2000) "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64: 413–28.
- Dalton, R.J. (2000) 'The Decline of Party Identifications', in Russel J. Dalton and Martin P. Wattenberg (eds) *Parties Without Partisans Political Change in Advanced Industrial Democracies*, pp. 19-37. Oxford: Oxford University Press.
- Dalton, R.J. and Wattenberg, M.P. (2000) 'Unthinkable Democracy Political Change in Advanced Industrial Democracies', in Russel J. Dalton and Martin P. Wattenberg (eds) *Parties Without Partisans Political Change in Advanced Industrial Democracies*, Oxford University Press, Oxford, pp.19 - 37.
- Dalton, R.J., Mc Allister, I., Wattenberg, M.P. (2000) 'The Consequences of Partisan Dealignment', in Russel J. Dalton and Martin P. Wattenberg (eds) *Parties Without Partisans Political Change in Advanced Industrial Democracies*, pp. 19-37. Oxford: Oxford University Press.
- Davern, Michael, Todd Rockwood, Randy Sherrod, and Stephen Campbell (2003) "Prepaid Monetary Incentives and Data Quality in Face-to-Face Interviews: Data From the 1996 Survey of Income and Program Participation Incentive Experiment." *Public Opinion Quarterly* 67(1):139-47.
- Deschouwer, K. and Nuytemans, M. (2005) 'De Stemtest van de VRT: een kijk in de keuken', *Samenleving&Politiek*, Vol. 12, pp.13 - 31.
- Dillman, Don (1978) *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.

- Dillman, Don (2007) *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.
- Downs, A. (1957) *An Economic Theory of Democracy*, Harper & Row, New York.
- Eady, G., Dufresne, Y., van der Linden, C., Loewen, P. and Hove J. (2012) 'Advancements in proximal measures for Voter Engagement Applications' (Working Paper).
- Eijk van der, C., van der Brug, W., Kroh, M., Franklin, M. (2006) 'Rethinking the dependent variable in voting behavior: On the measurement and analysis of electoral utilities'. *Electoral Studies* 25(3): 424-447.
- Erikson, R. S., Tedin, K. (2007) *American Public Opinion: Its Origin, Content, and Impact*. New York: Wiley.
- Evans, G., Heath, A.F. and Lalljee, M.G. (1996) 'Measuring left-right and libertarian-authoritarian values in the British electorate', *British Journal of Sociology*, Vol. 47 No.1, pp.93 - 112.
- Evans, J. (2004) *Voters and Voting – An Introduction*, Sage Publications, London.
- Farrell, D., Schmitt-Beck, R. (2008) *Non-Party Actors in Electoral Politics. The Role of Interest Groups and Independent Citizens in Contemporary Election Campaigns*. Baden Baden: Nomos.
- Fiorina, M.P. (1981) *Retrospective Voting in American National Elections*. New Haven, CT: Yale 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*, Vol. 2 No.4, Article 7.
- Fivaz, J. and Schwartz, D. (2007) 'Nailing the Pudding to the Wall – EDemocracy as Catalyst for Transparency and Accountability', presented at the *International Conference on Direct Democracy in Latin America*, Buenos Aires.
- Fivaz, J., Jeitzinger, B. (2005) 'E-Democracy Beyond E-Voting' in: P. Cunningham and M. Cunningham (eds), *Innovation and the Knowledge Economy. Issues, Applications*, pp. 590-597 Case Studies. Amsterdam: IOS Press.

Fivaz, J., Schwartz, D. (2007) 'Nailing the Pudding to the Wall – EDemocracy as Catalyst for Transparency and Accountability', presented at: International Conference on Direct Democracy in Latin America, Buenos Aires.

Franklin, M. (2003) *Voter Turnout and the dynamics of electoral competition*, Cambridge University Press, Cambridge/New York

Garzia, D. (2010) 'The Effects of VAAs on Users' Voting Behaviour: An Overview', in Cedroni, L and Garzia, D. (eds.), *Voting Advice Applications in Europe. The State of the Art*, ScriptaWeb, Napoli, pp.13 - 33.

Gemenis, K. (2011) 'Estimating the positions of political actors through voting advice applications: some methodological considerations', *Acta Politica* forthcoming. Available at: <http://www.utwente.nl/mb/pa/staff/gemenis/>

Gendall, Philip, Janet Hoek and Don Esslemenont. 1995. "The Effect of Appeal, Complexity and Tone in a Mail Survey Covering Letter". *Journal of Market Research Society* 37 (3): 251-68.

Griffiths, Thomas & Steyvers, Mark (2004). "Finding scientific topics". *Proceedings of the National Academy of Sciences* 101 (Suppl. 1): 5228–5235.

Groot, L.F.M. (2003) *De kwaliteit van stemprogramma's en de Stemvork*, SISWO, Amsterdam.

Hansen, Robert A. (1980) "A Self-Perception of the Effect of Monetary and Non-Monetary Incentives on Mail Survey Respondent Behavior." *Journal of Marketing Research* 17 (1): 77-83.

Harvey, Lee (1987) "Factors Affecting Response Rates to Mailed Questionnaires: A Comprehensive Literature Review." *Journal of Market Research Society* 29 (3): 341-353.

Held, D. (1996) *Models of Democracy (Second Edition)*. Cambridge: Polity Press.

Hooghe, L., Bakker, R., Brigeovich, A., De Vries, C., Edwards, E., Marks, G., Rovny, J. and Steenbergen, M. (2010) 'Reliability and Validity of Measuring Party Positions: The Chapel Hill Expert Surveys of 2002 and 2006', *European Journal of Political Research*, Vol. (4): 684-703,

Hooghe, M. and Teepe, W. (2007) 'Party profiles on the web: an analysis of the logfiles of nonpartisan interactive political internet sites in the 2003 and 2004 election campaigns in Belgium', *New Media & Society*, Vol. 9 No. 6, pp.965 - 985.

Hooghe, M., Teepe, W. (2007) 'Party profiles on the web: an analysis of the logfiles of non-partisan interactive political internet sites in the 2003 and 2004 election campaigns in Belgium'. *New Media & Society* 9(6): 965-985.

Inglehart, R. (1990) *Culture Shift in Advanced Industrial Society*. Princeton: Princeton University Press.

Inglehart, R., (1999) *Modernization and Postmodernization. Cultural, Economic and Political Change in 43 Societies*, Princeton University Press, Princeton.

Irwin, G.A., Holsteyn van, J.J.M. (2008) 'Scientific Progress, Educated Guesses or Speculation? On Some Old Predictions with Respect to Electoral Behaviour in the Netherlands'. *Acta Politica* 43: 180-202.

Iyengar, S., Simon, A.F. (2000) 'New Perspectives and Evidence on Political Communication and Campaign Effects'. *Annual Review of Psychology* 51(1): 149-169.

Jadot, A., Cautrès, B., Krouwel, A. and Vitiello, T. (2012) 'Designing a Voting Advice Application about a presidential election: issues or candidates? Insights from the French 2012 "Vote Compass"', paper presented at Advancing Electoral Research, ELECDDEM Final Conference 28-30 June Florence/EUI.

Jenkins R. (2006) 'Mass Media and Third-Party Insurgency' in Brady, H.E. and Johnston, R. (eds) *Capturing Campaign Effects*, pp. 356-382. Michigan: The University of Michigan Press.

Johnston, R., Andre B., Brady, H. E., Crete, J. (1992) *Letting the People Decide: Dynamics of a Canadian Election*. Kingston: McGill-Queen's Press.

Kies, R. and Dumont, P. (2011) 'Smartvote in Luxembourg: Usage and Impact of a New Electoral Tool', Paper presented at *the European Consortium for Political Research (ECPR) General Conference*, August 24-27 2011. Reykjavik, Iceland.

Kleinnijenhuis, J. & Krouwel, A.P.M. (2008). Simulation of decision rules for party advice websites. In F. Welsch, J.V. Carrasquero, A. Oropeza & C.B. Chen (Eds.), *Proceedings of the 2nd international multi-conference on society, cybernetics and informatics* (pp. 138-145). Orlando: IIS.

Kleinnijenhuis, J. en Krouwel, A.P.M. (2007) 'The Nature and Influence of Party Profiling Websites', Gepresenteerd op het politicologenetmaal 2007 te Antwerpen.

Kleinnijenhuis, J., Scholten, O., Atteveldt, W.H. van, Hoof, A.M.J. van, Krouwel, A.P.M., Oegema, D., Ridder, J.A. de, Ruigrok, P.C. & Takens, J. (2007). *Nederland vijfstromenland: De rol van de media en stemwijzers bij de verkiezingen van 2006*. Amsterdam: Bert Bakker.

Klingemann, H., Hofferbert, R.I., and Budge, I. (1994) *Parties, policies, and democracy*, Westview Press, Boulder (CO).

Klingemann, H., Volkens, A., Bara, J. , Budge, I and MacDonald, M. (2006) *Mapping Policy Preferences II. Estimates for Parties, Electors, and Governments in Eastern Europe, the European Union and the OECD, 1990-2003*, Oxford University Press, Oxford.

Krosnick, Jon A. 1999a. "Maximizing Measurement Quality: Principles of Good Questionnaire Design." In Robinson, J. P. , Shaver, P. R., and Wrightsman, L. S. (Eds.), *Measures of political attitudes*. New York: Academic Press, 37-57.

Krosnick, Jon A. 1999b. "Survey Research." *Annual Review of Psychology*, 50: 537-567.

Krosnick, Jon. A., Sowmya Narayan., and Wendy R. Smith. 1996. "Satisficing in Surveys: Initial Evidence." In Braverman, M. T., and Slater, J. K. (Eds.), *Advances in survey research*. San Francisco: Jossey-Bass, 29-44.

Krouwel, A (2012). *Party Transformations in European democracies*. Albany: SUNY Press (State University of New York Press).

Krouwel, A. (forthcoming) *Party Transformations in Advanced European Democratic States*, SUNY Press, New York

Krouwel, A. and Elfrinkhof, van A.M.E. (2009). From Text to Position: New Ways of Party Positioning on Political Issues and Dimensions. Paper presented at PoliticologenEtmaal, Bergen Dal 2009.

Krouwel, A. and Wall, M. (forthcoming) 'From text to political positions. Constructing visualisations of political systems from text', in Kaal, B., E. Maks and A. van Elfrinkhof (Eds). In preparation. From Text to Political Positions: Converging approaches to estimating party positions.

Krouwel, A., Fiers, Stefaan (2008). 'From Text to Position. Text-extraction and Authorisation of Party Positions along Dimensions of Political Competition', presented at: Voting Advice Applications between Charlatanism and Political Science, Antwerp.

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

Ladner, A., Fivaz, J., Joelle, P. (2010) 'Impact of Voting Advice Applications on Voters' Decision-Making', presented at: Internet, Politics, Policy 2010: An Impact Assessment, Oxford Internet Institute (OII), University of Oxford.

Laros, J. (2007) *Kiezen voor Stemadvies. Over de effecten van het gebruik van stemadviesprogramma's*, Unpublished Master Thesis, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.

Lau, R.R., Redlawsk, D.P. (2001) 'Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making'. *American Journal of Political Science* 45(4): 951-971.

Laver, M., and Garry, J. (2000) 'Estimating Policy Positions from Political Texts', *American Journal of Political Science*, Vol. 44 No.3, pp.619 – 34.

Laver, M., and Hunt, B. (1992) *Policy and Party Competition*, Routledge, New York and London.

Laver, M., Benoit, K. and Garry, J. (2003) 'Extracting Policy Positions from Political Texts Using Words as Data', *American Political Science Review*, Vol. 97 No.2, pp.311 - 331.

Linden, C. van der, Dufresne, Y., Hove, J., Eady, G and Loewen, P. (2012) Avoiding the Traps in VEA's. Pre-campaign Market Research and Micro-level Weighting Adjustments, paper presented at the panel: Dangerous Toys? Assessing the Electoral Integrity of Voter Engagement Application Websites, IPSA World Congress, Madrid, July 2012

Loosveldt, Geert, and Nathalie Sonck (2008) "An Evaluation of the Weighting Procedures for an Online Access Panel Survey." *Survey Research Methods* 2 (2): 93-105.

Louwerse, T. and Rosema, M. (2011) 'The Design Effects of Voting Advice Applications: Comparing Methods of Calculating Results', Paper presented at *the European Consortium for Political Research (ECPR) General Conference*, August 24-27 2011. Reykjavik, Iceland.

Lupia, A. (1994) 'Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections'. *The American Political Science Review* 88(1): 63-76.

Lupia, A., McCubbins, M.D. (1998). *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* . Cambridge: Cambridge University Press.

Marcus, G.E., Nueuman, R., MacKuen, M. (2000). *Affective Intelligence and Political Judgement*. Chicago: University of Chicago Press.

Marks, G., Hooghe, L., Steenbergen, M.R., Bakker, R., (2007) 'Crossvalidating data on party positioning on European integration', *Electoral Studies*, Vol. 26 No. 1, pp.23 - 38.

Marschall, S. (2005) 'Idee und Wirkung des Wahl-O-Mat'. *Aus Politik und Zeitgeschichte* 51-52: 41-46.

Marschall, S. and Schmidt, C.K. (2010) 'The Impact of Voting Indicators: The Case of the German Wahl-O-Mat', in Cederoni, L. and Garzia, D. (eds.), *Voting Advice Applications in Europe. The State of the Art*, ScriptaWeb, Napoli, pp.61 - 68.

Marschall, S., Schmidt, C.K. (2010) 'The Impact of Voting Indicators: The Case of the German Wahl-O-Mat', in: Cederoni, L., Garzia, D. (eds.), *Voting Advice Applications in Europe. The State of the Art (E-Book)*, pp. 61-68.

McDermott, M.L., (2006) 'Not for Members Only: Group Endorsements as Electoral Information Cues'. *Political Research Quarterly* 59(2): 249-257.

McGuire, W. J. (1986) 'The Myth of Mass Media Impact: Savagings and Salvagings'. *Public Communication and Behavior* 1:173-257.

Mikhailov, S., Laver, M. and Benoit, K. (2008) Coder reliability and misclassification in Comparative Manifesto Project Codings, Prepared for presentation at the 66th MPSA Annual National Conference, Palmer House Hilton Hotel and Towers, April 3–6, 2008.

Platek, Richard, and Carl-Erik Särndal (2001) "Can a Statistician Deliver?" *Journal of Official Statistics* 17, 1: 1-20.

Popkin, S.L. (1991) *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Praag, P. van (2007). *De stemwijzer: hulpmiddel voor de kiezers of instrument van*

Rabinowitz, G. and Macdonald, S. E., (1989) 'A Directional Theory of Issue Voting', *The American Political Science Review*, Vol. 83 No. 1, pp.93 - 121.

Rao, Kumar, Olena Kaminska, and Allan L. McCutcheon (2010) "Recruiting Probability Samples for a Multi-Mode Research Panel with Internet and Mail Components." *Public Opinion Quarterly*, 74(1): 68-84.

Redline, Cleo, Julia Oliver and Ron Fecso (2004) "The Effect of Cover Letter Appeals and Visual Design on Response Rates in a Government Mail Survey." Paper presented at the annual meeting of the American Association for Public Opinion Research.

Ruusuvirta, O. and Rosema, M. (2009) 'Do online vote selectors influence electoral participation and the direction of the vote?', presented at the *European Consortium for Political Research (ECPR) General Conference*, Potsdam.

Sartori, G., (1976) *Parties and Party Systems: A Framework for Analysis*, Cambridge University Press, New York.

Scheufele, D.A., Tewksbury, D. (2007) 'Framing, Agenda Setting, and Priming: The Evolution of Three Media Effects Models'. *Journal of Communication* 57(1): 9-20.

Sills, Stephen, and Chunyan Song (2002) "Innovations in Survey Research. An Application of Web-Based Surveys." *Social Science Computer Review* 20 (1): 22-30.

Thomassen, J., Aarts, K. and Van der Kolk, H. (2000) *Politieke Veranderingen in Nederland 1971-1998: Kiezers en de Smalle Marges van de Politiek*, SDU Uitgeverij Den Haag.

Traschel, A., Mair, P. (2009) 'When Parties (also) Position Themselves: An Introduction to the EU Profiler' Robert Schuman Centre for Advanced Studies Working Paper: EUI RSCAS; 2009/65; EUDO - European Union Democracy Observatory

Trechsel, A. and Mair, P. (2011) 'When Parties (also) Position Themselves: An Introduction to the EU Profiler', *Journal of Information Technology & Politics*, Vol. 8, pp.1 - 20.

Triga, V, Baka, A and Figgou, L. (2012) 'Neither agree, nor disagree': A critical analysis of the middle answer category in Voting Advice Applications. Paper prepared for the Workshop "VAA Research – State of the Art and Perspectives Schloss Mickeln, Düsseldorf, March 2-3, 2012

van der Brug, W., Pellikaan, H. (2003). 'Preface'. *Acta Politica* 38: 1-4.

Vavreck, L. (2001) 'The Reasoning Voter Meets the Strategic Candidate'. *American Politics Research* 29(5): 507-529.

Wagner, Markus and Ruusuvirta, Outi (2011) "Matching voters to parties: Voting advice applications and models of party choice." *Acta Politica* forthcoming. Available at: <http://www.palgrave-journals.com/ap/journal/vaop/ncurrent/index.html>

Walgrave, S., van Aelst, P., 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.

Walgrave, S., Nuytemans, M., Pepermans, K. (2009) 'Voting Aid Applications and the Effect of Statement Selection'. *West European Politics* 32(6): 1161 - 1180.

Wall, M. Krouwel, A. and Kleinnijenhuis, J. (2011) Comparing Voters and Parties Online: The Effects of Multi-Attribute Utility Decision Rule Section in Vote Advice Applications, paper presented at *PoliticologenEtmaal*, June 9-10 2011. Amsterdam, the Netherlands.

Wall, M., Krouwel A. and Kleinnijenhuis, J. (2012) Political Matchmakers – How do the decision rules employed by Vote Advice Application Sites influence their advice? Paper presented at ELECDem Conference, EUI June 28-30 Florence.

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*, forthcoming.

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*, Vol. 14 No.3, pp.203 - 218.

Zaller, J. (1992) *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

Zeileis, Achim, Christian Kleiber, and Simon Jackman. 2010. "Regression Models for Count Data in R." Online manual.