Aalto-yliopisto
Aalto-universitetet
Aalto University

Aalto University
School of Science
Master's Programme in Information Networks

Designing Voting Advice Applications

The Finnish Case

Veikko Isotalo

Supervisor: Professor Risto Sarvas

Advisor: Professor Jyrki Wallenius

Master's Thesis
Espoo, 18 November 2020

Aalto-yliopisto
Aalto-universitetet
Aalto University

Author Veikko Isotalo

Title of thesis Designing Voting Advice Applications: The Finnish Case

School School of Science

Programme Information Networks **Code** SCI3047

Thesis supervisor Prof. Risto Sarvas

Thesis advisor(s) Prof. Jyrki Wallenius

Date 18.11.2020 **Number of pages** 98 + 10 **Language** English

Abstract

Voting advice applications (VAAs) have become a part of electoral campaigns in many European countries. There is evidence that the usage of VAAs have real political consequences, as the usage has been linked to effects on political knowledge, turnout and vote choice. However, several methodological and normative issues have been pointed out regarding VAAs. The thesis follows design science research strategy, which identifies and proposes solutions to issues related to five Finnish voting advice applications that were developed for the 2019 parliamentary election. Based on the identified issues, a new overall VAA design is envisioned along with a novel way to visualize VAA results.

The research data consisted of news articles about the VAAs and data collected from the VAAs, including VAA statements and candidate responses. To identify issues directly from the VAAs, content analysis along with expert evaluation was performed. Additionally, an interactive web application was developed with R programming language to visualize VAA output.

The results showed that the Finnish VAAs lacked transparency, user interactivity and visualizations that would allow for candidate comparisons. Moreover, the VAAs also had algorithmic issues and problems with VAA statement structures. A novel network-based approach comparing candidate to candidate closeness was developed and it also proved to be suitable for detecting possible strategic behavior of parties in guiding their candidates' VAA answers. Based on these findings a new VAA design was envisioned, which consisted of two phases. The suggested design aimed to give users more power over VAA designers in the process and improve the quality of the voting advice.

It is left for future research to implement and test the suggested design. The goal of this thesis is to provide guidance to VAA developers and encourage for more innovative VAA designs.

Keywords voting advice application, candidate-based VAAs, VAA design, Finnish parliamentary elections

Aalto-yliopisto
Aalto-universitetet
Aalto University

Tekijä Veikko Isotalo

Työn nimi Vaalikoneiden suunnittelu: Tarkastelussa suomalaiset vaalikoneet

Korkeakoulu Perustieteiden korkeakoulu

Koulutusohjelma Informaatioverkostot **Koodi** SCI3047

Vastuuopettaja/valvoja Prof. Risto Sarvas

Työn ohjaaja(t) Prof. Jyrki Wallenius

Päivämäärä 18.11.2020 **Sivumäärä** 98 +10 **Kieli** Englanti

Tiivistelmä

Vaalikoneista on tullut osa poliittisia kampanjoita monissa Euroopan maissa. Vaalikoneiden käytöllä on todettu olevan poliittisia seuraamuksia, sillä vaalikoneiden käyttö on yhdistetty poliittiseen tietotasoon, äänestysaktiivisuuteen ja puoluevalintaan. Kuitenkin vaalikoneilla on havaittu olevan useita menetelmällisiä ja normatiivisia ongelmia. Tämä diplomityö hyödyntää design tieteeseen pohjautuvaa tutkimusstrategiaa, joka tunnistaa viiden 2019 eduskuntavaaleissa käytetyn suomalaisen vaalikoneen keskeiset ongelmat ja ehdottaa näihin ratkaisuja. Tunnistettujen ongelmien pohjalta kehitellään uusi malli vaalikoneelle ja uusi tapa esittää vaalikonetuloksia.

Tutkimusaineisto koostui vaalikoneisiin liittyvistä uutisartikkeleista ja vaalikoneista kerätyistä aineistoista, mukaan lukien vaalikoneväittämistä ja ehdokkaiden vastauksista. Vaalikoneiden ongelmien havainnointi tapahtui sisältöanalyysin ja asiantuntija-arvion avulla. Lisäksi toteutettiin vuorovaikutteinen web-sovellus R-ohjelmointikielellä vaalikoneen tulosten visualisointiin.

Tutkimustulokset paljastivat, että suomalaisilla vaalikoneilla oli puutteita läpinäkyvyyden, vuorovaikutteisuuden ja ehdokkaiden välisten vertailujen suhteen. Vaalikoneissa havaittiin myöskin algoritmeihin ja väittämiin liittyviä ongelmia. Lisäksi kehitettiin interaktiivinen verkostopohjainen sovellus, joka mahdollisti ehdokkaiden läheisyyden vertailun toisiin ehdokkaisiin. Sovelluksen avulla pystyi myös tunnistamaan puolueiden vaikutusyrityksiä ehdokkaidensa vastauksiin. Näiden tulosten pohjalta suunniteltiin uudenlainen vaalikone, joka koostui kahdesta vaiheesta. Tämän vaalikonemallin tavoitteena oli lisätä käyttäjien valtaa vaalikoneen suunnittelijoihin nähden ja parantaa vaalikoneen tarjoaman suosituksen laatua.

Tulevan tutkimuksen tehtäväksi jää toteuttaa ja testata suositeltua vaalikonemallia. Tämän diplomityön tavoite on neuvoa vaalikoneen kehittäjiä ja kannustaa uusiin vaalikoneinnovaatioihin.

Avainsanat vaalikoneet, ehdokaskeskeiset vaalikoneet, vaalikonemuotoilu, eduskuntavaalit

Contents

Preface

I want to thank my thesis advisor prof. Jyrki Wallenius for comments and my thesis supervisor prof. Risto Sarvas for making writing the thesis a lean process.

I also want to thank my partner for all the support and understanding which made writing of this thesis possible.

Espoo, 18 November 2020
Veikko Isotalo

Symbols and abbreviations

Symbols

$\boldsymbol{U}(\boldsymbol{V}, \boldsymbol{C})$	voter's $\boldsymbol{V}$ utility for candidate $\boldsymbol{C}$

Operators

$\|\boldsymbol{V}\|$	Euclidean length of a vector $\boldsymbol{V}$
Σ_i	sum over index i
$\boldsymbol{V} \cdot \boldsymbol{C}$	dot product of vectors $\boldsymbol{V}$ and $\boldsymbol{C}$

Abbreviations

DSR	design science research
EFA	exploratory factor analysis
HBL	Hufvudstadsbladet
HS	Helsingin Sanomat
MDS	multidimensional scaling
PCA	principal component analysis
SVT	Suomen virallinen tilasto
UI	user interface
VAA	voting advice application
Yle	Finnish Broadcasting Company

1 Introduction

Voting advice applications (VAAs) are online tools that help voters pick a candidate/party by comparing voters' preferences on issues to the stances of candidates/parties (Marschall and Garzia 2014, 1). VAAs work as cognitive short cuts to voters, as they help voters in selecting a candidate who to vote for. The cognitive short cut function of the VAA is especially important in the Finnish electoral context, where voters are facing an informational overload, as voters of populated districts need to select one candidate among hundreds of alternatives (see von Schoultz and Papageorgiou 2019). It is not a surprise that VAAs have become increasingly popular among the Finnish electorate, nearly half of the voters reported to have used them during the campaigning period for the 2019 Finnish parliamentary election (SVT 2019).

Internationally VAAs are a success story, as they have become a lasting feature of electoral campaigns in many European countries, such as the Netherlands, Germany, Switzerland, Denmark and Finland (see Marschall and Garzia 2014; Germann and Gemenis 2019). Voters are not the only ones who have set their eyes on VAAs, as researchers have identified the potential of using the VAA data to learn more about political elites' (and VAA users') stances on issues (see e.g., Isotalo et al. 2020; Tromborg 2019; von Schoultz and Papageorgiou 2019). Moreover, voting advice application usage has been linked to higher levels of reported turnout, also positive effects have been found on knowledge increase and vote choice (Munzert and Ramirez Ruiz 2020). However, more research is still needed to test the existence of these effects (Munzert and Ramirez Ruiz 2020).

Despite the popularity of the VAAs, they have not been without problems. For instance, there have been cases of strategic or manipulative behavior of candidates regarding statement responses to increase their visibility in the VAA results, also VAA algorithms that match users to candidates have not functioned as the developers expected. In addition to manipulation attempts and coding errors, VAA research has also identified methodological and normative issues that can erode the validity and reliability of the VAA recommendations (see e.g., Gemenis 2013; Fossen and van den Brink 2015). However, typical VAA users can be completely oblivious to these issues, if VAAs lack transparency. It is a common practice by the majority of the Finnish VAA developers not to disclose how users are matched with candidates. Yet, I consider discussion of VAA design to be more important than ever before, as VAA usage has become more common.

The main audience of this thesis are VAA developers and designers. In this thesis, I will identify and propose solutions to issues related to voting advice applications that were developed for the Finnish 2019 parliamentary election, following the design science research strategy. My research objective is to envision a VAA design that would improve the quality of the voting advice. To reach the research objective I have formulated two prerequisite research questions: (RQ1) "What are the main problems with most popular Finnish VAAs?" and (RQ2) "How should VAA results be displayed to voters for them to gain insights regarding candidates' positions to other candidates?". The main research question being (RQ3) "What kind of VAA design would solve identified issues with Finnish VAAs and provide voters more control over the VAA process?"

To answer the main research question, I will first represent the relevant literature regarding VAAs. Then I will inspect five different Finnish VAAs (Yle, Helsingin Sanomat, Alma Media, MTV Uutiset and HBL) by performing expert evaluation of the VAAs and reading news articles and documentation regarding the VAAs to identify issues related to them. After this I will present a visualization method that solves one of the issues identified with Finnish VAAs. Lastly, I will propose an overall VAA design that counters the detected problems and improves the quality of the voting advice.

To my knowledge, there is no VAA research that would have applied design science approach in identifying issues within particular VAAs and proposing a holistic VAA redesign that solves these issues. Although there is research on general methodological issues with VAAs (see e.g., Gemenis 2013) and studies that propose improvements in algorithmic design (e.g., Romero Moreno 2020; van der Linden and Dufresne 2017), none of the pre-existing studies consider the local context (electoral system, political landscape) and propose a VAA design that is custom-tailored to the setting in question.

The thesis is structured as follows: in chapter 1, I will explain what VAAs are, why they are important and the history of VAAs in Finland. Chapter 2 focuses on the existing literature and the theoretical background of the thesis. The chapter starts with inspecting VAA functionality from user and algorithmic perspectives, also revealing the political science theories behind VAA algorithms. The chapter is concluded by presenting methodological and normative critique of VAAs. Chapter 3 revisits the research objective and the research questions. Chapter 4 presents the research methodology along with data and methods. This is followed by results (chapter 5) that are presented in three parts, each part answering one research question. In chapter 6, the thesis is concluded by the discussion that points out limitations of the thesis and prospects for future research.

1.1 What are voting advice applications and why are they important?

The definition of a voting advice application according to Marschall and Garzia (2014, 1) states that VAAs are online tools assisting voters in their electoral decisions by comparing voters' policy positions with stances of parties and/or candidates. Policy positions are measured by filling a web-survey that compares user's answers to candidates' answers (Marchall and Garzia 2014, 1). After the comparison, the application produces an output which ranks parties (and/or candidates) by the level of agreement with the voter/user, additionally VAAs can display a graph that indicates user's position in policy dimensions in respect to parties (Marschall and Garzia 2014, 1).

Other definitions of VAAs also exist that do not require comparison of candidates' and user's answers, for instance Suojanen (2007) defines VAAs as mere information procurement tools regarding parties' and candidates' policy positions. Fossen and Anderson (2014) have problematized the current VAAs focus on matching voters to candidates, which inexplicitly supposes that the central issue of democratic process is citizens' ignorance about party positions. It suffices to say that VAAs do not need to be only about matching voters to candidates, but they could also be used to facilitate deliberation (see more Fossen and Anderson 2014). Here, I will limit this thesis to the "matching VAA" model following the definition of Marschall and Garzia (2014, 1), as this is the prevailing framework that most VAAs rely on.

There is no single ideal form of a VAA that would suit to all political contexts. However, in the *Lausanne declaration* (Garzia and Marschall 2014, 227–228) certain standards and minimal requirements have been established for all VAA makers. The general standards state that VAAs should be "open, transparent, impartial and methodologically sound" (Garzia and Marschall 2014, 227). The openness means that VAAs should be accessible for all voters. Transparency refers to transparency of intentions and funding regarding the VAA platform development, and it also means that users should know how parties and candidates are positioned on the issues. Moreover, transparency encompasses the matching algorithm, which should be publicly available and explained to the users. According to impartiality, all parties should be primarily included to the VAA and the VAA design should not favor any party in a systematic manner (Garzia and Marschall 2014, 227–228).

Talponen and Salminen (2007, 47) agree on these standards, as it is important that voters can trust the VAA's functioning to use it as a support tool for decision-making. Sound design,

meaningful content and good overall usability are especially important features in VAA development (Talponen and Salminen 2007, 48-49). Talponen and Salminen (2007, 46–47) also provide insight on different roles that voters, candidates and service providers have in providing the voting recommendation. Both voters and candidates input their personal opinions on issues to the VAA, but for voters the output is the voting advice, whereas for candidates VAA is essentially an election advertisement (Talponen and Salminen 2007, 46–47) with the distinction that it facilitates sharing of candidates' issue positions to the voters. The service provider maintains the website (input) and receives user traffic and user data (output), both of which provide financial opportunities. In figure 1, I will inspect the roles of these actors. I have further expanded an original figure from Suojanen (2007, 26), where VAA developers are synonymous to service providers. I, however, consider these two actors to be separate from each other, as developers can include people and organizations that do not work for the service provider (e.g., political scientists, web-application developers), even though it is to be expected that service providers (e.g., Helsingin Sanomat) have their own representatives in the development team. Therefore, I consider developers and service providers to be partially overlapping.

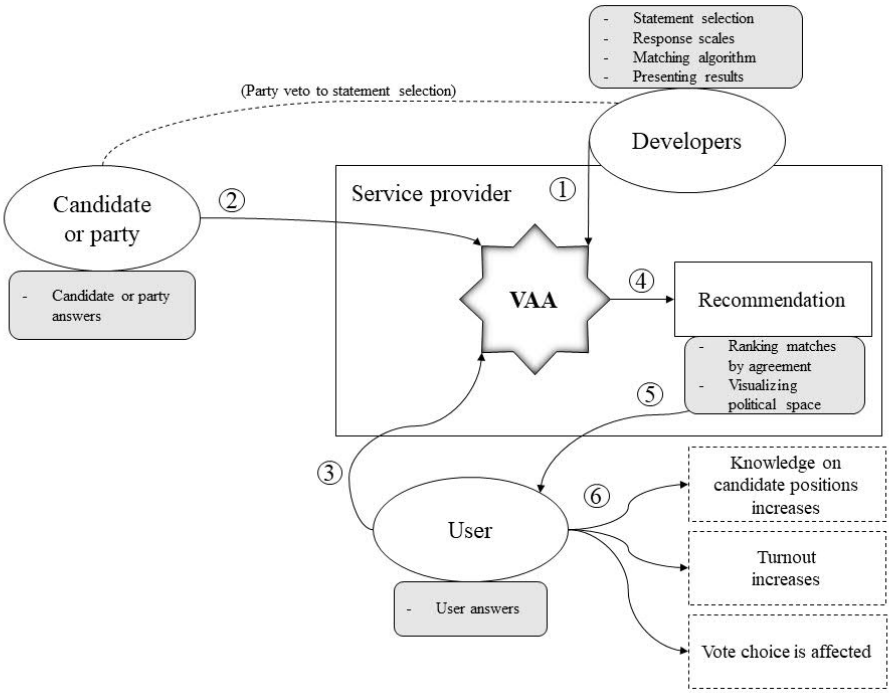

Figure 1. VAA actors and their roles in making the voting advice. Adopted and modified from Suojanen (2007).

Figure 1 displays developers, users and candidates interacting with the VAA. The VAA itself is located within the service provider. In the initial stage (1), the developers design the VAA.

This is a time-consuming process, as it requires the developers to select question-statements, choose appropriate response scales, implement the matching algorithm and determine how results should be presented for the users. In the second stage, candidates or parties are asked to submit their responses to selected VAA question-statements. Alternatively, developers can ask experts to fill parties' answers to secure truthful answers. In some VAAs, parties have been given the opportunity to veto statements, if they considered them to be unfair. After the second stage, the developers still have an opportunity to tweak and fine-tune the VAA before launching it. After the launch, users input their own answers and the VAA algorithm computes the recommendation tailored to their answers. The voting recommendation is presented to the voters usually by ranking candidates by agreement percentages and, in some VAAs, users are placed among candidates in low-dimensional ideological space. After presented with the voting advice, user's knowledge on candidate/party positions increases (Munzert et al. 2020; Kamoen et al. 2015; Paloheimo 2007). However, in a recent meta-analysis by Munzert and Ramirez Ruiz (2020) the effect of VAAs on citizens' knowledge on parties' issue positions was deemed to be rather small, although positive. Munzert and Ramirez Ruiz (2020) call for more evidence to confirm this effect.

VAAs' effects on turnout and vote choice are also conflicting, for example Andreadis and Wall (2014) found that using VAAs resulted in increased likelihood of vote switching, Garzia et al. (2017) have concluded that VAA users are more likely to vote than non-users, a finding that was further backed by Germann and Gemenis (2019) who found *smartvote* (a popular VAA in Switzerland) to be responsible of 1.2% of the turnout in the 2007 Swiss federal election. In contrast to these findings, an experimental design of Munzert et al. (2020) found no effects on either. The general picture is complicated, as Munzert and Ramirez Ruiz's (2020) meta-analysis reported strong positive effects of VAA usage on reported turnout and vote choice, but large heterogeneity in effect sizes is apparent mostly due to differences in study designs, as some studies have been plagued with self-selection issues. The existence of these issues in the VAA research have been identified (see e.g., Pianzola 2014) and suggestions to tackle these issues have been made (see Germann and Gemenis 2019).

It is important to further elaborate that there are multiple types of VAAs. VAAs can have identical user interfaces and algorithms, but they can still vary in terms of level of recommendation and positioning method. The recommendation level can be either candidate-based or party-based. In practice, candidate-based recommendation level means that a candidate is recommended for the user, whereas party-based recommendation gives advice on parties. The majority of the VAA research has focused on party-based VAAs, as most

European VAAs (*StemWijzer*, *Wahl-O-Mat*, *VoteMatch*, *HelpMeVote*, etc.) provide recommendations regarding parties (see e.g., Gemenis and van Ham 2014; Dumont et al. 2014, 146). According to Dumont et al. (2014), candidate-based VAAs have only been in use in Switzerland, Luxembourg, Finland, Lithuania and Denmark for legislative elections. Dumont et al. (2014) note that main reason why these countries have employed candidate-based VAAs is that the countries in question employ preferential voting (i.e. permitting/coercing votes to be casted for individual candidates). Providing candidate-based recommendations can also be justified, if politics is very personalized, which means that individual candidates have high electoral significance in comparison to their parties. As mentioned before, it is evident that the level of recommendation is dependent on the underlying electoral system in which the VAA is embedded. There is also a possibility to have a hybrid of the two levels where a user is recommended both a party and a candidate. However, typically these hybrid recommendations are based on candidate-based recommendations that are transformed into party-level recommendations (e.g., calculating party average or median responses) without actual party-level response data. Therefore, I consider two recommendation levels to be sufficient for describing currently existing VAAs.

The second component determining the type of the VAA is the positioning method (see e.g., Gemenis and van Ham 2014). Positioning method can be either self-reporting or expert evaluation. Self-reporting is the simpler of the two methods, as it directly collects the parties'/candidates' positions on issues and the VAA uses these positions to match voters with parties/candidates. Expert evaluations provide an alternative for positioning parties on issues. These evaluations typically rely on public documents (e.g., party manifestos, law proposals and speeches) that indicate what policies parties/candidates support. Choosing between self-reporting and expert evaluations is a trade-off regarding susceptibility to manipulation. On one hand, self-reporting can be problematic, as candidates/parties can choose positions that they do not actually support to improve their image and electability. On the other hand, expert evaluations are also prone to errors and in some cases, experts do not agree on party positions which can result in inaccuracies, which would not have been present, if parties had been asked directly.

In figure 2, recommendation level and positioning method are used to construct a matrix, which lists all four types of VAAs. Of the four types candidate-based & expert evaluations VAAs are noted to be hypothetical, as this type of a VAA would be practically impossible to implement in the Finnish context, as experts would need to position thousands of candidates, some of which are making their entry to politics and do not yet have public documents

from which to infer their positions on issues. The most common VAAs are the party-based VAAs, however, in Finland candidate-based and self-reporting VAAs have been most prevalent. This is also going to be the type of VAAs which will be the focus of this thesis. Next, I will discuss the importance of VAAs for democracy.

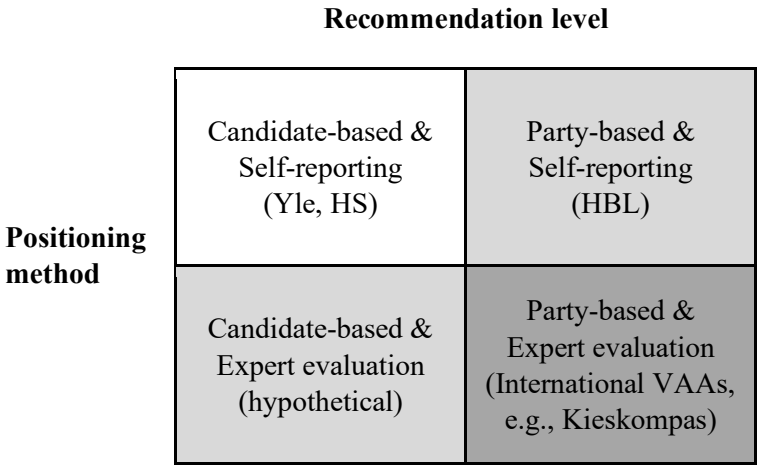

Figure 2. Four types of VAAs classified by positioning method and level of recommendation.

According to Anderson and Fossen (2014), VAAs are intended to strengthen democratic process by 1) informing voters about party positions on issues (citizen competence), 2) increasing voter turnout (political participation) and 3) ensuring that political representatives' attitudes match the electorate's attitudes (representation). Anderson and Fossen (2014) point out that VAAs seek to bridge "citizen competence gap", which reflects citizens' ignorance of candidate/party positions, by providing them with easily accessible information regarding political actors' standings on issues. In this view, VAA becomes a heuristic device, i.e. a short cut, that cuts down the costs of voting by finding a best-matching candidate from several possible candidates (Gemenis and Rosema 2014, 282). Making the vote choice easier is especially important in complex electoral contexts that overload citizens with information (e.g., in Uusimaa electoral district in Finland voters must choose from nearly 500 candidates). Voters tend to be more responsive to simple cues in complex settings that allow them to reduce time and effort in making the vote choice (Lau and Redlawsk 2006). The consequence of easing the vote choice can then lead to increase in voter turnout. However, the increase of turnout having a positive effect on democracy depends on which voter groups increase their electoral participation. If VAAs simply increase the turnout among highly educated and politically knowledgeable, while not mobilizing politically disenfranchised and voters with fewer resources, then the increase of turnout might not have a positive outcome

in terms of electoral representation. Therefore, it is important that one inspects who the VAA users are.

Typical VAA users have been young politically knowledgeable and highly educated males (see e.g., van de Pol et al. 2014; Marschall 2014). However, in the official statistics, VAA usage has become more common in all age groups in Finland, as 49 per cent of the electorate reported to have used VAAs prior to the 2019 parliamentary election (SVT 2019). On a positive note, Gemenis and Rosema (2014) found that VAAs increased electoral participation among young voters and voters less interested in politics in the 2006 Dutch parliamentary election. This result would suggest that VAAs have positive effects on the functioning of the democracy.

van de Pol et al. (2014) have further categorized VAA users based on their VAA usage intentions to doubters, seekers and checkers. Doubters are voters who have high uncertainty about their vote choice, and they have low political interest and external political efficacy (i.e., a view that government cannot respond to individuals' needs and demands) (van de Pol et al. 2014). Seekers are also uncertain about their vote and have little political interest, while having clearer party preferences and they are seeking guidance from the VAA (van de Pol et al. 2014). Checkers are politically interested and efficacious and they filled the VAA more for entertainment purposes, as they are quite certain which party they will vote for before filling the VAA (Van de Pol et al. 2014). In the *Kieskompas* self-reported dataset, majority of the users were checkers, roughly third were seekers and one tenth were doubters (van de Pol et al. 2014). In second order elections, i.e. subnational (municipal elections) and supranational (European Parliament), the share of doubters and seekers in the VAA users increases, as voters perceive that less is at stake in these elections (van de Pol et al. 2019). However, VAA users self-reported to be more interested in campaigning in second-order elections than in national elections, an unexpected finding which van de Pol et al. (2019) explain by scarcity of political information available that requires voters to invest more time and effort to become informed.

In terms of enhancing issue representation of voters, VAAs should provide unbiased and neutral recommendations that reflect issues on which voters base their vote choice. The inherent idea behind VAAs recommending candidates to voters, that have highest possible issue congruence, is that political representatives should be mirroring citizens' views on issues. This is called the delegate model of representation (Anderson and Fossen 2014, 223–225). Therefore, the main focus regarding representation has been on selecting VAA questions that capture relevant aspects of the political landscape. One should remember that the

effects of VAA usage are dependent on users' pre-existing preferences. It is suggested by Andreadis and Wall (2014, 127) that VAA usage can 1) strengthen user's pre-selected vote choice, 2) undermine user's pre-selected vote choice, or 3) both strengthen and undermine depending on the voter interpretation of the VAA recommendation output.

To conclude, VAAs are considered to be important tools for voters that cut costs of information retrieval regarding political candidates/parties, thus possibly increasing turnout. VAAs also provide recommendations that promote voting for candidates that match with user's views on issues. VAAs have real-life political consequences, as they are not merely tools for entertainment. It is no wonder that VAAs have been credited to enhance democracy. According to Anderson and Fossen (2014, 226), the underlying model of democracy, political participation and representation, that the VAAs are currently based on, are largely contested concepts of "social choice theory of democracy; a minimalistic voting-centered conception of political participation; and a delegate model of democratic representation". These issues will be discussed also in chapter 2. Next, I will present the history of VAAs in Finland.

1.2 VAAs in Finland

The first VAA in Finland was developed by Yleisradio (the Finnish Broadcasting Company [Yle]) for the 1996 European Parliament election (Vihtonen 2007). The 1996 Yle VAA, is among first VAAs in the world, even though offline VAAs had existed already in the 1980s Netherlands (see Marschall and Garzia 2014, 2). Erkki Vihtonen, a project manager in Yle, had adopted and further developed the idea of a VAA from CNN that had a quiz on their website that determined which presidential candidate was more suitable for the user (Suojanen 2007, 15). Ever since the 1996 Yle VAA, in Yle VAAs, candidates have answered the VAA questions instead of using journalists or experts to position them on the issues (Suojanen 2007, 15). According to Suojanen (2007, 15), the 1996 VAA did not amass more than 8000 users. The breakthrough of VAAs in Finland for a wider audience happened in the 1999 parliamentary elections, as 83 000 users used the Yle VAA and 68 per cent of candidates had created VAA profiles (Vihtonen 2001 in Suojanen 2007). In 1999, Helsingin Sanomat (a Finnish newspaper with highest circulation) launched their first VAA for the 1999 European parliament election (Suojanen 2007, 15). Following this development, the proliferation of Finnish VAAs happened in the early 2000s, as other media outlets (e.g., local newspapers and television broadcasting companies), interest groups and political parties developed their own VAAs. The number of VAAs in Finland peaked at 11 in 2003, after which

the number of VAAs has decreased and the most of the VAA usage has concentrated on Yle, Helsingin Sanomat and MTV Uutiset VAAs (Suojanen 2007).

The share of voters using VAAs in Finland has increased substantially from 23% in 2003 (Paloheimo 2007) to 49% in 2019 (SVT 2019). One factor behind the substantial increase of VAA usage is increase in Internet accessibility. In line with other research on VAA users, Paloheimo (2007) found that typical VAA users were politically knowledgeable, highly educated and young. The age gap in VAA usage still persists in 2019, as nearly 68% of voters in the age group 25–34 reported to have used VAAs prior the election, in contrast to 30% of 65–74-year-olds (SVT 2019). The influence of VAA recommendations on party choice was considered to be lower than on candidate choice among the electorate (SVT 2019). VAAs had either high or moderate influence on party choice for 38 per cent of the VAA users, whereas for the candidate choice the corresponding number was 53 (SVT 2019). In 2003 roughly one third of VAA users considered the VAAs to have either moderate or high influence on their candidate choice. This means that there has been a substantial increase in applying VAA output in voting related decision-making over the last 16 years. Moreover, the influence of VAA recommendations on candidate choice were stronger among younger voters than older ones (SVT 2019). This means that VAAs play more important role among young voters' information retrieval and party/candidate choice. However, it should be noted that young voters tend to also seek candidate related information online (e.g., via social media, party websites, candidate blogs and newspaper websites) more actively than older age groups (SVT 2019).

2 Literature review

In this chapter, I will inspect the functionality of the VAAs from theoretical and design perspectives. The chapter will be concluded by introducing critique of VAAs presented in existing literature. Firstly, I make a distinction between user interface (UI) and server-side. The user interface consists of elements with which users interact, whereas the server-side computes the output (i.e., the voting advice) based on user inputs (e.g., numeric answers to set of questions) following a certain algorithm. I will be looking into both user interface and server-side elements, the focus of the chapter being on the latter. It should be noted that I will not discuss implementation of VAAs in terms of web technologies, as the scope of this thesis is limited to conceptual and design choices in making VAAs.

In 2.1, I will present typical user interface elements of VAAs (e.g., questions, response scales and result visualizations) that users interact with. Next, I will focus on the side of the VAAs that is typically hidden from the users, namely the server-side of the VAA. Constructing a voting recommendation from a set of policy question answers is not a trivial task. In 2.2, I will present the underlying political science theories on which VAAs are built upon. In 2.3, I will explore the algorithmic designs of VAAs and explore different technical solutions that have significant consequences regarding the voting recommendation. In 2.4, methodological and normative critiques of VAAs are discussed, which also concludes this chapter.

2.1 The anatomy of a VAA

The anatomy of a VAA refers here to the user interface elements of the VAA. These elements are presented in order of appearance, as user navigates through the VAA in figure 3. First, before answering the VAA questions, the user typically chooses their location in the landing page. The location determines which electoral district's candidates are matched with the user. However, party-based VAAs do not require location data of the user, as parties' answers are the same in all districts.

In the second step, the user engages with the VAA questions. Questions can be revealed to the user all at once or one by one. If all questions are revealed at once, some Finnish VAAs provide live match tracking, which means that the users see a list of best matching candidates that is being updated as they progress in filling the VAA statements. There is currently no research on the topic how this affects and possibly biases users' answers.

The VAA questions, also known as issue statements or question-statements (all three terms are used interchangeably), are the most important element of the VAA. The quality of the VAA, referring to validity and relevance, is directly dependent on the questions that it uses. Walgrave et al. (2009) have shown that VAA statement selection, i.e. selecting certain configurations of VAA statements, has a profound impact on voting recommendations. Therefore, the VAA questions need to be carefully selected and they should exhibit some of the following characteristics: 1) relevant for the upcoming election, 2) the questions and their response options are easy to understand, 3) capture inter- and intra-party differences between parties and candidates, 4) connected to the agenda of the electoral body's next term and 5) provoke emotional responses in voters or concern important economic issues (Haukio and Suojanen 2004). In practice, VAA questions can take multiple forms, e.g., there can be completely hypothetical questions that are disconnected from day-to-day politics, but they are suitable for VAAs, because they provoke emotional responses and candidates are divided on those issues. Alternatively, there are questions that voters might not be that much emotionally invested in, but they are in the agenda of the next parliamentary term, and therefore, they have political relevance. VAA questions can also be categorized by theme (e.g., economy, immigration, health care) or by the underlying ideological dimension (e.g., Left–Right, Conservative–Liberal). VAA designers typically seek to have a balanced set of questions so that multiple political themes and ideological dimensions are represented.

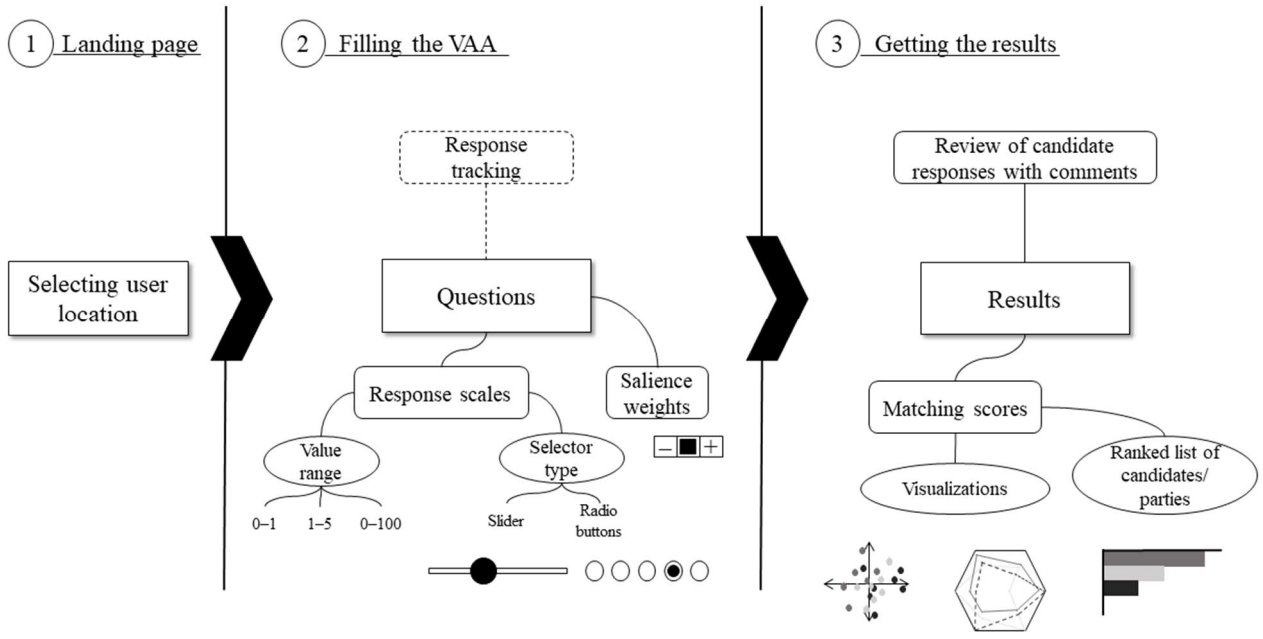

Figure 3. UI elements in different phases of using the voting advice application.

Users interact with the questions via response scales. Response scales can be numerical or categorical (Suojanen 2007, 26). Nominal response scales are rarely used, as they do not contain information about rankings of response categories. Typically, users and candidates are asked to either agree or disagree with a statement or selecting a value between these two extremes. Numerical and ordinal response scales can easily handle such response options. Common value ranges in numerical response scales are "yes/no" (binary, 0–1), "yes/no + don't know", "yes/neutral/no + don't know" and 5-point Likert scale (1–5, where 1 represents "completely disagree" and 5 "completely agree") (see Wagner and Ruusuvirta 2012). Also, other much wider scales exist (e.g., 0–100); however, these are not widely used. Likert scales have been popular in Finnish VAAs and internationally (Suojanen 2007, 26; Gemenis 2013, 270). According to Gemenis (2013, 270), the central idea of Likert scale is that individual items are imperfect measurements of latent variables, and Likert scales seek to capture these latent variables, as responses from multiple items are combined. In contrast to Likert scale, interval scales are used in many expert and mass surveys to measure the respondent's position on the latent variable directly (Gemenis 2013, 270). For example, in Mendez (2017), 11-point interval scales are used to measure subjective evaluations of respondents' own positions on Left–Right dimension. Gemenis (2013, 271) also adds that the advantage of Likert scale is that it measures political actors' and users' positions with same items and scales, thus claiming to provide voting advice to voters.

Response scales can be displayed to users as radio buttons or sliders. Sliders are preferred, if the value range of the response scale is high, but responses indicated with sliders can be difficult to replicate, if they are sensitive. Rosema and Louwerse (2016) have shown that varying the length of response scales has a significant impact on VAA recommendations, meaning that different scales provided different voting advice for the same users. Rosema and Louwerse (2016) tested only for mechanical effects that were detected via post-response recoding of the scales, but the authors also suggest the possibility of psychological effects.

In addition to positioning themselves on issues, the VAA users and candidates can, in some VAAs, determine salience weights (i.e. how important the issue is). This should be asked separately, as Likert scales cannot measure salience (Gemenis 2013, 270), meaning that it is possible that extreme position on an issue is also connected to high salience, but it does not have to be. Salience weights are quite common in VAAs, but not all VAAs have them, as it has been pointed out by VAA providers that many users do not use them (Wagner

and Ruusuvirta 2012, 409). Salience weights can be binary, or they can have multiple levels (see Wagner and Ruusuvirta 2012).

In the final step, users are provided with the results of the VAA. The results consist of the voting advice, which is represented by user's matching scores with candidates and parties. Usually matching scores are presented in percentages from 0 to 100% (zero meaning complete disagreement and hundred complete agreement). Mendez (2017, 50) has also introduced a matching score scale from –100 to 100, where zero is the middle-point of the scale representing neutral stance or equal amount of disagreement and agreement with the candidate.

The candidates/parties are typically ranked in a descending order by matching scores. Thus, users are shown a list of candidates, not just a single candidate. The list is typically the central element of the vote advice. In addition to the list, after seeing the results, users can review individual candidates' responses in greater detail and read candidates' comments explaining their choice of issue positions. These comments are useful for users and researchers when trying to understand candidates' motives (see e.g., Ylä-Anttila 2014). These textual answers can possibly influence users' views on issues and even increase users' knowledge on issues and not just candidate positions. However, there is currently no research regarding the effects of text answers in VAAs.

The voting advice application output also usually contains visualizations of candidate and user positions. In some cases, visualizations provide the main voting advice, especially, if lists of candidates with matching scores are not provided. However, this can be problematic, as Bruinsma (2020) has found that users presented with identical visualizations of VAA results have low levels of agreement on what party is closest to their positions. The visualizations in VAAs seek to present either high-dimensional or low-dimensional data (Bruinsma 2020). The dimensionality of the visualizations depends on the design structure of the VAA: if matching users to candidates happens in low-dimensional space (e.g., one to five dimensions), then the visualizations are also low-dimensional. High-dimensional visualizations typically utilize bar plots, whereas low-dimensional visualizations utilize spider graphs or spatial maps (Bruinsma 2020). According to Bruinsma (2020), none of the visualizations are perfect, but spider graphs tend to be most disliked by users. This could be because they have multiple problems in terms of interpretability (see Bruinsma 2020).

Spatial maps are usually two-dimensional and give users an indication of their own positions and the political actors' positions in the political space (Germann et al. 2015). Users can then compare their positions to candidates'/parties' positions via two different distance

metrics: Euclidean and Manhattan (Bruinsma 2020). These distance metrics will be explained in detail in the following subchapters. Fossen and van den Brink (2015) have pointed out that the VAA designers' selection of political axes, which are represented in the VAAs, is a political act.

Now, that the UI elements of the VAA have been presented, I will shift my focus on the server-side calculations of the VAA. The server-side calculations are responsible for providing the voting advice. Next, I will introduce the political science theories that act as the basis of the VAA algorithms. These theories highlight the importance of issues in politics and thus justify the existence of VAAs.

2.2 Political science approaches on issue convergence of political actors and voters

The usefulness of the VAA stems from the idea that voters pick their candidates based on political issues. This issue-centered view of politics has long traditions in political science (see e.g., Downs 1957). The most influential theory explaining how issues relate to electoral outcomes is the spatial model of issue voting, according to which issue positions of voters and candidates determine vote choices of voters. There are also other theories which emphasize the capabilities of candidates to promote issues in which they have an advantage and increase their salience.

Political science does not, however, claim that issues are all that matters in politics, as party identification (Campbell et al. 1960), socio-demographic factors (Lipset and Rokkan 1967) and personal vote earning attributes (Shugart et al. 2005) have been found to be influential on explaining candidate choice of voters. As a matter of fact, in open-list proportional representation systems (such as Finland) the likelihood of voters to vote according to candidate–voter issue proximity decreases (Singh 2010), which is due to the complex electoral environment, where voters have hundreds of potential candidates to vote for in a single district. Lau and Redlawsk (2006) note that voters tend to resort to simple cues, also known as heuristics, to reduce the informational overload in choosing the right candidate. To counter these tendencies, VAAs seek to bring issues back to politics and serve as a tool to connect voters to candidates by serving as an issue-emphasizing short cut for voters. Next, I will present three models of issue voting that have been linked to the VAAs by Wagner and Ruusuvirta (2012).

2.2.1 Proximity and Directional models

There are two main approaches on how voters choose the candidate to vote for within the framework of spatial models of issue voting, namely the proximity model (Downs 1957) and the directional model (Matthews 1979; Rabinowitz and Macdonald 1989). A lot of effort has been put into demonstrating which model is the most consistent with the actual voting patterns of voters (e.g., Westholm 1997; Macdonald et al. 1998). Evidence for both models exist, but Lewis and King (2000) note that the underlying statistical assumptions involved with applying these models have not been properly tested. Merrill and Grofman (1999) suggest that these two models can be combined into a single unified model that applies logics of both models. Merrill and Grofman (1999) have tested their model on two European multi-party systems and concluded that the mixed proximity and directional model is the best predictor of actual party positions. Grofman (1985) has also suggested that some voters discount the policy positions of candidates, meaning that voters evaluate candidates' policy positions in relation to the current status quo by applying a discounting factor. The discounting model can take a form of a proximity or a directional model or a mixture of the two depending on the value of the discounting factor (see Merrill and Grofman 1999, 8–9). However, more recent research has been more supportive of the proximity model over the directional model (Lacy and Paolino 2010), many studies still highlight the value of both models, as voters might be implied to apply directional logic to challenger party candidates and proximity to incumbents and candidates of which policy positions they are more aware of (Cho and Endersby 2003). Thaler (2015) notes that the directional logic can be thought of as a form of emotional thinking, whereas the proximity logic highlights a mode of technical thinking. Kropko and Banda (2018) note that the framing of an issue affects whether voters apply directional or proximity voting and both logics are present among the electorate. Thus, both models are deemed important and can be used in the context of VAAs. Next, these models are presented and later explored in detail with one- and two-dimensional examples.

The proximity model, also known as the Downsian proximity model, was originally introduced in the 1950s by Anthony Downs (Downs 1957), after which it has been further developed and formalized for multi-dimensional and multi-party scenarios. The proximity model is a form of a spatial model. Assumptions underlying spatial models according to Rabinowitz and Macdonald (1989, 93) are "1) each voter can be represented by a point in some hypothetical space such that reflects the person's ideal set of policies, 2) the policy position of

each candidate can be represented by a point in the same space". The proximity model further assumes that voters vote for the closest candidate to their policy position, hence the word "proximity". The original Downsian model was presented as a part of a more ambitious undertaking called the economic theory of democracy (Downs 1957), which is considered as one of the foundations of rational choice theories of voting (Evans 2004).

The proximity model employs Hotelling's (1990) spatial model designed to explain rational consumer behavior in economics. Where Hotelling measures consumers' (demand) physical distance to businesses (supply) on a single dimensional line, Downs replaces the physical distance with ideological distance, rational consumers with rational voters and the businesses with candidates/parties. Instead of focusing on plethora of political issues Downs' argues that voters use ideological positions as short cuts to identify candidates that maximize their utility. This follows due to uncertainty of the voters, as voters are not perfectly informed on all issues, candidate positions and policy outcomes, moreover, obtaining candidate positions on all political issues would be costly and time-consuming. The original Downsian model applies a single ideological dimension that represents the economic Left–Right dimension. If voters are evenly or normally distributed along the dimension it follows that in a two-party system parties should place themselves to the median voter position to maximize their vote (Downs 1957). This is also known as the median voter theorem, which represents a stable equilibrium of party positions in two-party systems (Downs 1957). Therefore, it becomes evident that the proximity model has important implications to candidate strategies, where candidates should position themselves.

The utility, i.e. voter's evaluation of the candidate, in the proximity model is a declining function of policy distance from the voter to the candidate (Merrill and Grofman 1999). The utility is maximized when the voter and the candidate have identical spatial position (Merrill and Grofman 1999). The functional form the utility function can take a quadratic or a linear form, the difference of the two being that in the quadratic proximity utility the distance between the voter and the candidate is squared (Merrill and Grofman 1999). Here we specify the utility with the quadratic version following Merrill and Grofman (1999):

$$U(\boldsymbol{V}, \boldsymbol{C}) = -\sum_{i=1}^{n} (v_i - c_i)^2$$

where $\boldsymbol{V} = (v_1, \dots, v_n)$ and $\boldsymbol{C} = (c_1, \dots, c_n)$ are voter and candidate vectors in n-dimensional policy space, v_i and c_i are voter and candidate positions in the ith dimension, $U(\boldsymbol{V}, \boldsymbol{C})$

represents the voter's $\boldsymbol{V}$ utility for candidate $\boldsymbol{C}$, and a negative sign in the formula indicates that the utility decreases with distance (Merrill and Grofman 1999).

The directional model is in contrast with the proximity model. There are multiple directional models, but here I focus on two of them, namely models from Matthews (1979) and Rabinowitz and Macdonald (1989). The directional model stems from the theory of symbolic politics, where an issue conveys a symbol and this symbol may provoke an emotional response due to associations attached to the symbol (Rabinowitz and Macdonald 1989, 94). Rabinowitz and Macdonald (1989) claim that this is the natural way for people to respond to political issues, instead of conceptualizing them in spatial terms. For an issue to have political weight, it must provoke a symbolic response, which has two quantities: a direction and an intensity (Rabinowitz and Macdonald 1989, 94). The direction of the response can be either favorable or unfavorable towards the issue, whereas intensity refers to the strength of the emotional response regarding the issue (Rabinowitz and Macdonald 1989, 94). Of the two elements mentioned above, Matthews (1979) model takes account only the direction component. Matthews utility function can be written as follows:

$$U(\boldsymbol{V}, \boldsymbol{C}) = \frac{\boldsymbol{V} \cdot \boldsymbol{C}}{\|\boldsymbol{V}\| \|\boldsymbol{C}\|} = cos\theta$$

where, $\boldsymbol{V} \cdot \boldsymbol{C}$ are a scalar product of the two vectors, $\boldsymbol{V}$ and $\boldsymbol{C}$ represent voter and candidate vectors, θ is the angle between $\boldsymbol{V}$ and $\boldsymbol{C}$, $\|\boldsymbol{V}\|$ and $\|\boldsymbol{C}\|$ are Euclidian lengths of the vectors (Merrill and Grofman 1999).

Rabinowitz and Macdonald (1989) model, RM model for short, takes in the account also the intensity. In the RM model, voter's utility for a candidate is a scalar product of the voter and candidate vectors (Merrill and Grofman 1999):

$$U(\boldsymbol{V}, \boldsymbol{C}) = \boldsymbol{V} \cdot \boldsymbol{C} = \sum_{i=1}^{n} v_i c_i$$

where, $\boldsymbol{V}$ is a vector for a voter and $\boldsymbol{C}$ for a candidate.

The consequence of the RM utility function is that the RM model favors extreme candidates, as the utility grows without limit while the candidate places themselves further from the neutral point of the dimension. This conclusion can be contested, as there is no empirical evidence that would show that voters always pick the most extreme candidates over moderate candidates. If it were so, democratic countries would have more extreme left- or right-wing governments in place. To counter this weakness in the model Rabinowitz and

Macdonald (1989) augment the model by adding an area of acceptability, which effectively leaves out extreme candidates that are positioned too far off the neutral point. In a two-dimensional space the area of acceptability would be a circle around the neutral center point. Rabinowitz and Macdonald (1989) are vague about the implementation of the acceptability constraint, as there appears to be no clear rule on how it should be applied. Iversen (1994) criticizes RM model's area of acceptability, as it is unlikely that voters would have a single standard to evaluate all candidates' acceptability, and proposes that each voter should have their own area of acceptability (e.g., for an extremely conservative voter a moderately liberal candidate might be out of the question). Including original formulation of the area of acceptability, the optimal strategy for candidates according to RM directional model would be to pursue aggressive strategies, but still to remain within the area of acceptability (Rabinowitz and Macdonald 1989, 109–110). Next, the models are explored via examples.

First, to illustrate how the proximity and the directional model function I will present two one-dimensional examples. The first example represents an election with three candidates located on a single dimension (see figure 4). We can call this dimension as the Left–Right, where negative values represent leftist positions and positive values rightist ones. The dimension ranges from -4 to +4, where 0 position represents the middle point (center) of the dimension. Next, we can inspect a vote choice of a single moderately rightist voter located in +1. According to the proximity model, the voter should pick the closest candidate to their position. In this case, there are two candidates C1 and C2 that are equally distant from the voter. The utilities of C1 $(-(1-(-1))^2)$ and C2 $(-(1-3)^2)$ are -4, and the candidate C3 has the highest distance to the voter, its utility being -9. As there are two equally close candidates to the voter, the voter is indifferent between these candidates (C1 and C2), as their expected utility for the voter is the same. Therefore, the voter might pick either of the two at random. This case highlights the fact that according to the proximity model voter does not pay attention to the direction of the candidate (left or right).

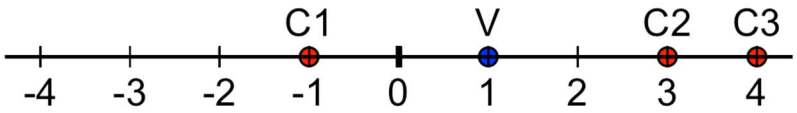

Figure 4. Candidate choice in one-dimensional spatial space.

If we apply the RM directional model on the same voter and candidate locations, and the neutral position locates at 0, the voter would prefer the candidate C3, because the utility

function of the RM directional model is simply a scalar product of the candidate and voter positions. For the candidate C1 the utility is negative ($1 \times (-1) = -1$), for C2 it is 3 ($1 \times 3 = 3$) and for C3 it is 4 ($1 \times 4 = 4$). Alternatively, we can apply the Matthews directional model in which the voter assigns the highest utility to candidates C2 and C3, whereas the candidate C1 has a negative utility. This is so, because the Matthews directional model is only interested in the angle between the candidate and the voter position, and in a one-dimensional model there are only two directions.

Next, two two-dimensional examples will be presented. In the first one, I will calculate utilities of each candidate for a single voter. In the second one, differences between models are shown graphically by drawing voters' indifference lines. In figure 5, there is a single voter in a two-dimensional space (Left–Right and Liberal–Conservative), both dimensions ranging from -4 to +4, where 0 is the neutral point. The voter is in position (1,4) which is slightly on the right and extremely conservative. There are two candidates C1 (4,1), who is extremely rightist and moderately conservative, and C2 (1,-1) who is identically rightist as the voter, but slightly liberal on the second dimension. According to the proximity model and applying the quadratic utility function, the voter should pick candidate C1, as the utility for C1 is -18 ($-(\sqrt{(1-4)^2 + (4-1)^2})^2$), and for C2 -25 ($-(\sqrt{(1-1)^2 + (4-(-1))^2})^2$). The voter–candidate distances are drawn to the left panel of figure 5. In other two panels of figure 5, I apply directional models to the same situation. Matthews directional model is only interested in the angle between voter and candidate position vectors, whereas RM model includes the vector length to the calculation. The utility in the RM model for candidate C1 is 8 ($1 \times 4 + 4 \times 1$) and for candidate C2 ($1 \times 1 + 4 \times (-1)$) is -3. Matthews model divides the scalar product by the vector lengths. Thus, the utility for C1 is 0.47 (which translates in angle of 62 degrees) and for C2 -0.69 (133° angle). Thus, according to the directional model the voter prefers candidate C1, as the candidate is on the same side as the voter in both dimensions.

In figure 6, indifference lines are drawn between two candidates. These lines present the positions where voters are indifferent towards the two candidates (C1 and C2). If the voter locates anywhere else in the policy space, the voter prefers either of the two candidates. In the upper left panel of the figure, I have placed the Matthews directional model's indifference line which intersects the neutral point of the two dimensions forming a 45-degree angle between the indifference line and the Left–Right axis. Intensity of candidate positions do not matter in this model, only the direction, thus Matthews model perceives C2 to be located in (1,0), hence the 45° angle. In the RM model (located in the upper right corner), the

indifference line also passes through the neutral point, but taking account the intensity of the candidate positions it is perpendicular to a line that connects the candidate positions. In the proximity model (lower left corner), the line of indifference simply passes through the middle point of the line between the two candidates in a 90° angle. In the final panel all lines are shown together.

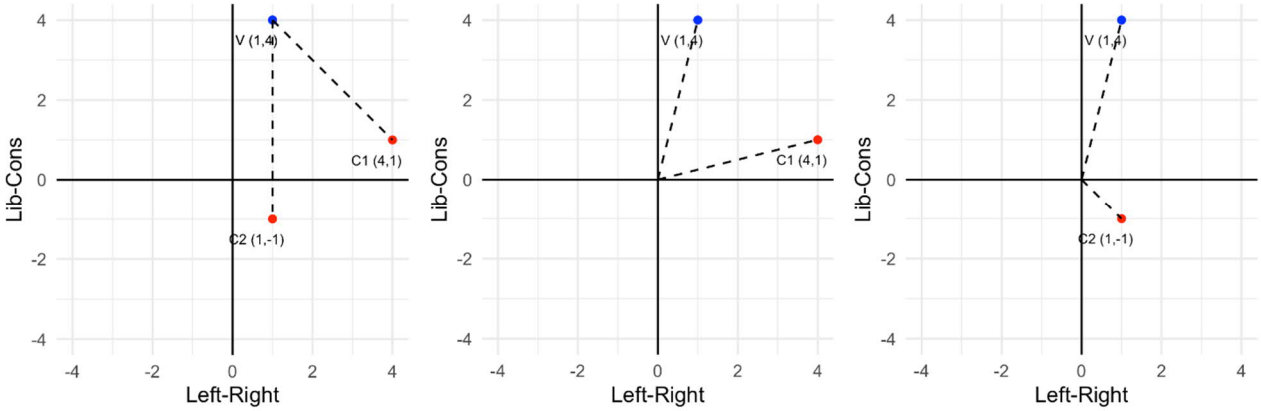

Figure 5. Candidate selection in two-dimensional spatial space.

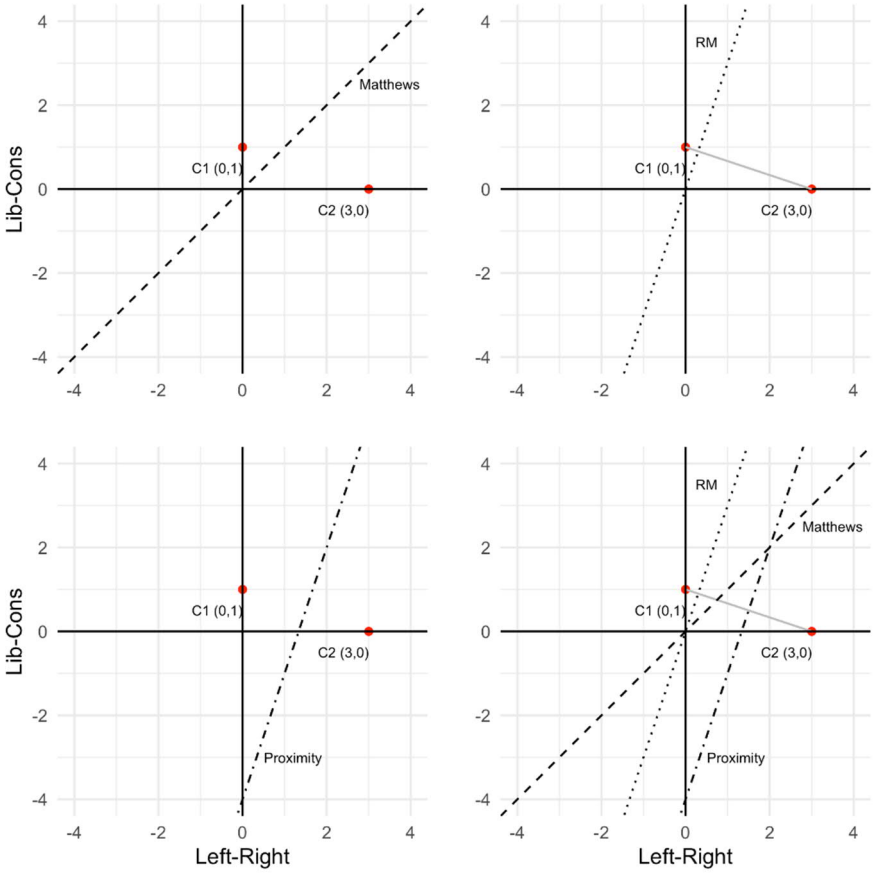

Figure 6. Lines of indifference in three models.

In the context of the VAAs, both the directional and proximity models can be implemented. According to Wagner and Ruusuvirta (2012), West European VAAs rely more heavily on the proximity model in their implementation, when matching VAA users to candidates/parties. Thus, focusing on the closeness of the candidate to the voter. This does not, however, mean that the directional model could not be applied to VAAs, as response alternatives to VAA questions are usually directional, meaning that the endpoints of the response scales are yes/no or agree/disagree. Also, it is not evident to which extent many VAAs follow the proximity model. Are VAA questions aggregated into larger issue dimensions, or is each question treated as an individual dimension? Are distances in multiple dimensions calculated with Euclidian distance, or is some other distance metric applied? These questions are discussed more in detail in 2.3, where I discuss the technical solutions that build up the VAA algorithm. The goal of this subchapter was to show how the spatial models of voting work, as these models are the foundations of issue-centered view of politics, a view that the VAAs are built upon. Also, it was shown that there are multiple ways to match voters to candidates according to these theories. Next, an alternative view of issue-centered politics is introduced that does not follow the spatial model, namely issue salience, which is also influential in the VAA design.

2.2.2 Issue salience

Issue salience refers to the importance of an issue for an individual voter or for the public in general. According to Budge and Farlie (1983), parties compete by focusing on issues on which they have a perceived advantage. More salient an issue, the more valuable it is for the party to own it. This strand of theorizing is also known as the issue ownership theory. It claims that parties have a differing reputation that is usually earned in handling political issues. Issue ownerships tend to change slowly, and they are formed by long-term actions of parties (Hosch-Dayican et al. 2013, 5). Slow shifts of issue ownership have been explained by links between political conflict and social structure (Petrocik 1996, 827). In addition to reputation, parties can also claim issue ownership via mere association that is a spontaneous response of the voters to link a party with an issue (Hosch-Dayican et al. 2013, 4). In sum, in the eyes of the voters, parties have become experts of taking care of certain issues which gives their candidates credibility on those issues.

As voters' issue ownership perceptions are quite fixed, candidates should shift their campaigning effort in maximizing the exposure of issues in which they have relative advantage

to other candidates (Petrocik 1996). This means that candidates should not pursue to change the minds of voters, but to highlight problems reflecting the issues on which they have an advantage (Petrocik 1996, 828). If salient issues of an election are successfully determined by some party, candidates of other parties are facing an uphill task to challenge the issue-dominant candidate(s).

One would expect candidates to avoid direct confrontation on policy standings with other candidates, if they are not owning the issue in question (Sigelman and Buell 2004). As mentioned previously, according to the issue ownership theory, the central objective in running a political campaign is to increase issue salience of one's own issues and avoid taking clear positions on issues that can decrease the support, in order to maximize the electoral base. However, empirical evidence from the US presidential elections points out that opposing candidates have been discussing on the same issues more than one would expect (Sigelman and Buell 2004). The authors account for the unexpected convergence of candidates in the same issues by highlighting the role of media pressuring candidates to talk about issues that are on the media's agenda (Sigelman and Buell 2004).

The VAAs serve exactly this function, by forcing the candidates to take a stand on issues, some of which might not be included in their campaigns, thus illuminating candidate positions and facilitating political debate. The difficulty for the VAA designer is to select appropriate salient issues on the VAA for an ideologically heterogenous electorate. Individual voters can have widely different preferences regarding the importance of political issues (e.g., green party voters may perceive environmental issues as the most important ones, whereas right-wing party supporters may evaluate the state of the economy to be the most important one). One solution by the VAA designers to respond to the individuality of voter issue rankings is to allow VAA users to apply weights to questions according to their own preferences. These single or multi-level weighting schemes have been more a rule than an exception in VAAs used in the Western European countries (Wagner and Ruusuvirta 2012, 404–405). Applying the weighting scheme, however, can make the recommendations of the VAA more difficult to understand for an average user. It has also been noted that many users do not use weights even though the option is available (Wagner and Ruusuvirta 2012, 409). Application of weights to candidate–user issue distances (e.g., doubling candidate–user issue distance, if a voter rates the issue to be twice more important than other issues) effectively implies using salience as a part of a standard proximity model (Wagner and Ruusuvirta 2012, 409).

Salience theory, which highlights the role of issues to election results, can also be implemented as a stand-alone method in the VAAs. Thus, the primary focus of a VAA would be on comparing agendas and issue priorities of voters and candidates. In this type of a VAA, the best matching candidate would not be a candidate that is closest to the voters on issue stances, but on what issues are the most important (Wagner and Ruusuvirta 2012, 408–409). Alternatively, one could think a Likert scale as a measure for salience, as voters who completely agree or disagree could also view the question to be more important for them than for voters that place themselves close to the neutral answer option. However, this is not advised, as salience and intensity of one's opinion are not identical concepts and Likert scales measure only intensity. Therefore, it would be optimal to measure salience and issue positions/intensity separately.

2.3 Algorithmic design in matching voters with candidates

In this part of the thesis, I will present the current state of the VAA algorithm design. I will identify five different types of algorithms, basing on the literature. I explore the properties of these algorithms by a fixed set of criteria: 1) identifying the data that is used for the recommendation (source of recommendation), 2) does the algorithm follow the issue-voting paradigm which has been the normative rationale behind the VAAs, 3) does the algorithm allow candidate and/or party matching, 4) the transparency of the recommendation, i.e. is the recommendation process understandable and how easy would it be for a user to reconstruct their recommendation scores, 5) the vulnerability to inside manipulation, i.e. if a party or a candidate answers untruthfully and finally 6) the vulnerability to outside manipulation, i.e. other attempts of manipulating the recommendations, excluding parties' and candidates' own attempts. After presenting the algorithm types, I will examine most important algorithmic elements (e.g., distance metrics and dimensionality) that are applied in these algorithms.

2.3.1 Types of VAA algorithms

The number of experimental VAA algorithms has increased significantly in the past 10 years. Most of this work is stemming from academic fields outside political science. Thus, it is no surprise that these algorithms do not stay true for the spatial model of issue voting. Yet, candidate–voter issue distance algorithms are still the most prevalent ones applied. Next, I

will present the variety of VAA algorithms and discuss their properties, strengths and weaknesses.

Table 1. VAA algorithm types

Algorithm	Issue distance	Issue distance extended	Learning VAA	Community-based recommendations	Candidate clustering and classification
Literature	E.g. Louwerse and Rosema (2014)	Pajala et al. (2018); Mendez (2017); Terán and Mancera (2017)	Romero Moreno et al. (2020)	Katakis et al. (2014); Djouvas and Tsapatsoulis (2018)	See Katakis et al. (2014); Terán and Mancera (2017)
Source of recommendation	User-candidate distance	User-candidate distance + other	Learning from users (or candidates) + user-candidate distance	Learning from users	Learning from candidates
Follows issue-voting paradigm	Yes	Partially	Yes	No	No
Candidate and/or party matching	Both	Both	Both*	Party only	Party only
Transparency of recommendation	Very high	Mediocre	Mediocre	Low**	Low
Vulnerability to inside manipulation	High/Low	High/Mediocre/Low	Mediocre/Low	Low	Mediocre
Vulnerability to outside manipulation	Low	Mediocre/Low	Mediocre/Low	High	Low

Note: * Learning VAAs can be used also for candidate-level recommendations, if one uses distance matrices and weights trained by matching candidates to their own parties. ** Transparency of the community-based algorithms is low in majority of the matching algorithms used.

In table 1, I have assigned VAA algorithms into five distinct categories that can be compared by their differing properties. The first of these is the traditional issue distance algorithm also known as "VAA 1.0", a term used by Terán and Mancera (2017) which refers to the static nature of the information, meaning fixed answers to set of questions, used for the voting recommendation. The issue distance algorithms can take many forms, i.e. these algorithms may use different distance metrics and varying number of dimensions that build up the recommendation (see Louwerse and Rousema 2014), but they all share the same properties regarding the source of the recommendation which utilizes solely candidate and voter

answers to VAA questions by calculating their distances. Naturally, issue distance algorithms stay true to the issue-voting theory and these algorithms make it possible to match users to either parties or candidates (depending on the VAA). In issue distance algorithms, candidate positions can be transformed into party positions by calculating candidate averages or median positions, or alternatively using party leader answers as indicative of the party positions, or simply asking from the party office for the party's official stances. These algorithms do not tend to be very complex and users may manually reconstruct the recommendation scores, if they know how distances are calculated, therefore, the transparency of the recommendation is labelled to be very high. In case of inside manipulation, candidates might not be truthful in their own answers, or parties might try to influence their candidates' answers by preparing them to answer in a certain way. This is an inherent problem with asking candidates directly about their views. In some VAAs, expert opinions are used to limit untruthful behavior of the parties, though utilizing expert opinions is not viable in candidate-centered VAAs, as is the case in the Finnish VAAs with over 2000 candidates per parliamentary election. Therefore, vulnerability to inside manipulation is either high or low depending on whether political actors are asked directly or not. As the data used for the recommendation is static in nature (candidate answers are collected before the launch of the VAA), there is a low risk of outside manipulation.

The second type of VAA algorithm is an extended version of the issue distance algorithm. In addition to utilizing user–candidate issue distances, these algorithms include other factors into the recommendation. Pajala et al. (2018) incorporated candidates' social media activity, newspaper mentions and political power (operationalized by incumbency status), Mendez (2017) utilized socio-demographic data of the user to enhance the recommendation accuracy while testing different set of modeling strategies for matching voters to parties, whereas Terán and Mancera (2017) included sentiment analysis vectors of candidates' social media posts and voters' perceptions of candidates to issues asked in the VAA. These additional factors are typically claimed to improve the recommendation quality while still holding on to the issue-voting paradigm. Terán and Mancera (2017) highlight that the inclusion of external data allows the creation of dynamic candidate profiles which they call "VAA 2.0", thus candidates are not matched with users based on merely static candidate profiles.

The additional factors can be related to individual candidates (e.g., Twitter activity of a candidate) or parties (e.g., sentiments toward issues in party manifestos) or VAA users themselves (e.g., gender or education level). Therefore, these types of algorithms can be used for both candidate and party matching. The additional factors in the extended issue distance

algorithms can cause issues with transparency of the recommendation, as these variables tend to complicate the recommendation process and, at worst, bias the recommendation results. In the case of Pajala et al. (2018), the inclusion of incumbency and traditional media variables into the recommendation algorithm gave new candidates a disadvantage, as the algorithm decreased their recommendation scores, thus making VAAs an uphill battle for inexperienced candidates. This contradicts the original idea of the VAAs, in which VAAs serve an equalizing function in terms of campaigning resource disparities to match voters with like-minded candidates. Therefore, one should be careful in including variables that could bias the recommendations of the VAAs to serve the political status quo and powerful political elites. The vulnerability to inside manipulation is labelled to range from high to low, as the labelling depends greatly on what type of external data is used in the recommendation. In a high vulnerability scenario, candidate's own tweets might be prone to manipulation, if the candidate knows that this data is collected for the recommendation algorithm, conversely sentiment of candidate's own tweets can have a balancing effect to untruthful VAA answers lowering the manipulation level to mediocre. On the other hand, chances of inside manipulation can be labelled as low, if expert evaluations are used to position parties on issues and objective measures unrelated to candidates' actions (e.g., voting records in parliament) are used as additional factors. Vulnerability to outside manipulation varies based on the external data from mediocre to low, as in the mediocre scenario the external data is user-created information (e.g., evaluations of the candidate) and these are prone to manipulation, whereas the VAA question section (static profile) of the recommendation is secure. In the low outside vulnerability scenario, the external data is some objective metric that cannot be easily manipulated.

The third category of VAA algorithms is called *learning VAA*, a concept launched by Romero Moreno et al. (2020), which presents a new algorithmic design that applies machine learning in the first phase to estimate two important parameters of the algorithm: distance matrices and salience weights linked to VAA questions. Distance matrices determine which issue-voting logic is applied to each question, ranging from proximity to directional logic or a combination of the two (see Mendez 2017). Salience weights determine the overall importance of the question in making the recommendation. In the second phase of the algorithm, distance matrices and weights are employed to calculate recommendation scores via traditional issue distance logic. The critical stage of the algorithm is the learning phase when distance matrices and weights are estimated. Romero Moreno et al. (2020) use self-identified issue-voters' voting intention to train their model by estimating the distance and weight

parameters so the parameters predict test-set issue-voters' voting intentions with the highest accuracy. It is, however, not mandatory to use voters to estimate the important parameters. Alternatively, in a Finnish electoral context with over 2000 candidates per parliamentary election, one could try to classify candidates to their own parties and estimate the important parameters of the model. To summarize, the learning VAA learns from users' answers and then applies user–candidate distance calculations which produce the final recommendation, thus the algorithm follows the issue-voting paradigm. Learning VAA can be used for matching both users to parties and users to candidates. However, for candidate matching to work there is a precondition that party classification is used to train the parameters which are then applied in user–candidate distance calculations. The transparency of the recommendation is classified to be mediocre, as users can theoretically calculate the recommendation scores, if they are presented with the distance matrices and salience weights. This way the users can be presented the logic behind the recommendation and see how the numbers constituting the recommendation score are formed. However, this is a rather complex task.

An important advantage of the learning VAA over the basic issue distance algorithm is that the learning VAA allows differing spatial voting logics to be applied while calculating user–candidate distances and that these logics are derived directly from the data (Romero Moreno et al. 2020). In addition, distance matrices also reveal how VAA questions are perceived by users/candidates which provide important information for researchers and VAA designers alike. A drawback of this approach is that it can also lead to nonsensical distance matrices, which can be indicative of a problematic model or that the question is not working as intended. I categorized both types of vulnerability of the algorithm to be either mediocre or low, because candidates do not know in advance whether a proximity or directional logic is applied to a specific question, which lowers the incentives for candidates to purposefully modify their answers. On the other hand, if users' answers are used for estimating the distance matrices, then there is a higher chance to manipulate the logic applied to a certain question, but candidate–user distances themselves still will not be affected. If candidates' answers are used to estimate the distance matrices, then there are no chances of external manipulation, as these distances are trained to predict these candidates' own party labels, which are objective facts, in comparison to users' subjective assessment of voting intentions of a certain party.

Community-based recommendation, also known as social vote recommendation, algorithms are similar to learning VAA algorithm in respect that they both utilize users' answers

to create recommendations. However, the community-based recommendations do not require candidates to answer VAA questions at all, as the recommendations rely solely on other users' answers and voting intentions (see Katakis et al. 2014; Djouvas and Tsapatsoulis 2018). Thus, the issue-voting paradigm is not applied by this algorithm. Community-based algorithm perceives the VAAs as one form of recommendation system. According to Katakis et al. (2014), these systems apply one of the two approaches in providing the recommendations: 1) content-based (i.e., suggesting similar items that the user has preferred in the past), or 2) collaborative-filtering (i.e., suggesting items that other users with similar preferences have liked). As VAA users' previous voting habits and party preferences are not known neither of the two approaches can be directly applied to VAAs. However, collaborative-filtering approach can be adjusted for the task of matching VAA users to parties, by calculating similarity of user profiles and inspecting users' k most similar users (Katakis et al. 2014). The voting advice of the VAA is simply the voting intention of these similar users. There is also a variety of different techniques that can be used to calculate community-based recommendations and most of them act like "black boxes" which lack transparency (Romero Moreno et al. 2020, 8). However, there are some exceptions, such as average voter method which has a very high transparency for the recommendation, as it simply replaces candidates with average voters of parties and calculates distances between the user and the average voters.

Katakis et al. (2014) found that community-based recommendations have higher accuracy in giving recommendations that match users' voting intention than standard issue distance methods between candidates and voters. Additionally, Tsapatsoulis et al. (2015) found out that voters perceived community-based suggestions to be more useful than traditional issue-based recommendations in an experiment where the recommendation method was not revealed to the voters. One possible explanation why the community-based recommendations have been more accurate than standard methods is that not all VAA users are issue-voters, meaning that other factors (such as party leader image and party identification) are more predictive of the vote choice than issue positions. The community-based algorithms try to match the user with their own community (or tribe) and individual users are expected to follow the recommendations of these communities. The algorithm supports a view of politics, where voters are divided into competing political tribes. This can be harmful to democracy due to possible effects of increasing polarization, thus eroding possibilities for parties and voters to collaborate and find compromise, even though they might agree on some issues. There are also other technical problems with community-based recommendations including a "cold start problem" which means that the recommendation engine requires a

sufficient number of initial users before the algorithm provides stable recommendations (see more Katakis et al. 2014). The requirement for initial users increases, if the number of parties is increased. It is also apparent that this type of algorithm cannot provide candidate-level recommendations in the Finnish context. The vulnerability of the algorithm to inside manipulation is low, as candidates do not have a say in the recommendations. The vulnerability to outside manipulation is much higher, further depending on whether there is a selection/verification process in place which filters out suspicious users that are used in training the algorithm.

The last type of algorithms used in VAAs is labelled as *candidate clustering and classification algorithms*. These algorithms include a variety of supervised and unsupervised machine learning methods such as k-means clustering, fuzzy c-means clustering, k-nearest neighbors, support vector machines and decision trees (see methods in Katakis et al. 2014; Terán and Mancera 2017). All these methods can be used in community-based recommendations, with the distinction that the algorithms are trained with user data. It is apparent that candidate clustering and classification does not follow issue-voting paradigm. In contrast, these algorithms perceive VAA recommendations as machine learning tasks. One significant drawback of these algorithms is that they are not viable options for making candidate-level recommendations, if the number of candidates is as high as in the Finnish elections. Moreover, these algorithms require a large number of candidates to provide party-level recommendations, which can reduce the applicability of these algorithms in electoral contexts that have only a few candidates running for an election. Candidate clustering and classification algorithms also have severe issues with transparency and interpretability, it being possible that spurious rules of classification exist and are undetected (Romero Moreno et al. 2020, 9). Moreover, the technical complexity of these models might be one reason why they are not more widely used (Romero Moreno et al. 2020, 9). Candidate clustering and classification algorithms can be susceptible to inside manipulation, as candidates can be untruthful in their answers, but this type of behavior can be detected by the algorithm, if also other candidates have similar answering patterns. These algorithms are not vulnerable to outside manipulation, as only candidates' answers are used as training data. Now that the existing VAA algorithms have been introduced, I will discuss design choices of algorithmic elements that can have significant effects for the recommendations.

2.3.2 Design choices of algorithmic elements

After choosing the algorithm for the voting advice application, the VAA designers must make five important design choices regarding the implementation of the VAA algorithm. The design choices include 1) choosing the number of dimensions of the modelling space, 2) choosing how questions are combined to form the dimensions, 3) choosing a distance matrix for an appropriate spatial model, 4) choosing a method to aggregate the overall distance scores between a user and a candidate in the modelling space and 5) choosing whether questions can be weighted. All five choices are non-trivial and different combinations of these elements can lead to significantly different recommendations.

The first design choice, namely choosing the number of dimensions for the modelling space, has no single standard practice. The number of dimensions can range from 1 to n in which n is the number of questions in the VAA. Typically, there are either low- or high-dimensional VAA models. For example, in a VAA comprising of 30 questions, a low-dimensional model would contain circa 1–5 dimensions, whereas a high-dimensional model would be closer to the total number of questions (30). In a low-dimensional model, these dimensions are no longer single-issue dimensions, but they represent multiple issues. One example of such a political dimension would be the Left–Right dimension, also called as "super issue" or "meta dimension" due to its capability to be linked to variety of issues (Inglehart and Klingemann 1976, 244). There are also some VAAs that have applied low to mid-range number of dimensions, for instance *EU Profiler* has used seven dimensions (Trechsel and Mair 2011). In this type of a setup, the dimensions reflect relatively wide issues that are linked to multiple statements in the VAA, e.g. types of issues could be immigration, environment and social policy. In a high-dimensional setting, individual dimensions represent directly individual issue statements in the VAA. In practice, both one-dimensional and high-dimensional models are typically unrealistic. One-dimensional model underestimates the presence of other latent factors, whereas high-dimensional model overlooks correlations among questions.

The second design choice, regarding how questions are combined to form the dimensions is also important. The simplest way to combine questions into dimensions is to sum them together, this implies the equal importance of all the questions for the dimensions. There are also three commonly used dimension reduction methods: 1) principal component analysis (PCA), 2) factor analysis and 3) multidimensional scaling (MDS). PCA uses uncorrelated linear combinations of original data items to construct principal components which seek to

preserve the variability of the data by maximizing the variance (Jolliffe and Cadima 2016). A significant drawback of the PCA is that it is a mere dimension reduction method, thus ignoring the underlaying structure in the data (see Costello and Osborne 2005). In contrast to PCA, the factor analysis (here we focus on exploratory factor analysis, also known as EFA) suggests that there are latent factors causing responses to individual items to covary and it aims to reveal them (Costello and Osborne 2005). Fabrigar and Wegener (2011) state that construct identification is one of the main uses of factor analysis, in case of VAAs this would be identifying latent political dimensions underlying VAA items. One should note that not all question items have equal importance in constructing the dimensions using the EFA, as some question items load more heavily on certain factors (dimensions), whereas some items might not load well at all. This results in some questions having more weight when candidates are positioned on the created factors. There are also multiple ways to calculate these factor scores (see DiStefano et al. 2009). van der Linden and Dufresne (2017) have suggested to use factor analysis inductively with a pilot survey to construct and validate ideological dimensions to be used in VAAs.

Multidimensional scaling has been used to construct low-dimensional approximations of distances between parties' (or users') VAA answers, as this method can be applied for a small number of data points (see König and Nyhuis 2020; Wagner and Ruusuvirta 2012; Louwerse and Rosema 2014). If VAA data is collected only from parties, then MDS might be the only viable option, as PCA and EFA require sufficient amount of observations to work. If PCA and EFA can both be used, then factor analysis (EFA) should be prioritized, as VAA questions measure opinions on issues and these opinions are influenced by underlying positions on latent dimensions (e.g., Left–Right, Conservative–Liberal and Authoritarian–Libertarian).

The third design choice is about choosing an appropriate distance matrix for determining user–candidate distances. The distance matrix implies which issue-voting logic is applied in the VAA (Mendez 2017). As it has been mentioned before there are two main logics: proximity and directional logic. Additionally, there is also a hybrid matrix that combines both logics used in *EUvox* voting advice application (Mendez 2017). The three matrices are visualized in figure 7. The values in the matrices represent distance scores associated with combinations of user and candidate answers on 5-point Likert scale ranging from completely agree (CA) to completely disagree (CD). Having a positive score of +1 means total agreement between the candidate and the user, whereas -1 represent total disagreement. It should be noted that this type of operationalization does not allow usage of Euclidean distance metric.

Different issue-voting logics can be applied to individual questions (this is how learning VAA algorithm functions), but this might be confusing for users. It could be possible to alleviate the user confusion by letting the users choose which logics are applied to which questions. The combinations of the first and the third design choices, regarding the number of dimensions in the VAA and application of different spatial models have been researched by Mendez (2017). Mendez (2017) concluded that proximity model performed better than directional model in low dimensions to predict voters' voting intention, whereas a directional model was more accurate in high dimensions.

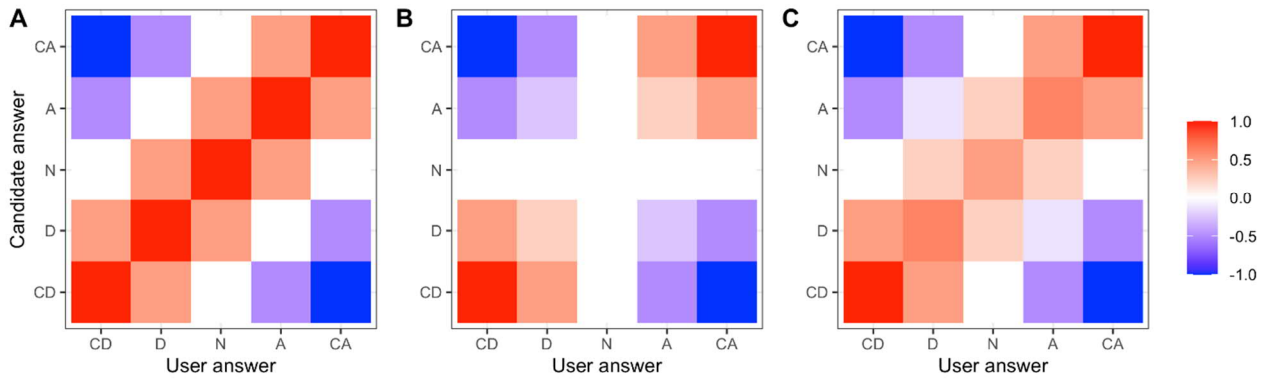

Figure 7. Distance matrices. From left to right proximity (Manhattan distance), directional (Scalar Product metric) and hybrid matrices (combination of the two). Adopted from Romero Moreno et al. (2020) and Mendez (2017).

The fourth design choice concerns the distance calculations. There are three main metrics for aggregating distance scores in multi-dimensional space: 1) Manhattan distance, 2) Euclidean distance and 3) Mahalanobis distance. Manhattan distance is typically used in high-dimensional VAA space calculations (Romero Moreno et al. 2020, 5). In calculating Manhattan distance, the distance scores of individual questions are simply summed together which is why the metric is also known as the city-block distance. Manhattan distance is calculated with the equation:

$$d(\boldsymbol{V}, \boldsymbol{C}) = \|\boldsymbol{V} - \boldsymbol{C}\| = \sum_{i=1}^{n} |v_i - c_i|$$

where $d(\boldsymbol{V}, \boldsymbol{C})$ denotes the distance, $\boldsymbol{V}$ is the voter and $\boldsymbol{C}$ the candidate vector in an n-dimensional space, v_i is voter's and c_i candidate's position on the ith dimension.

Euclidean distance is more often used in low-dimensional space (Romero Moreno et al. 2020, 3). Euclidean distances are straight-line distances between two points and these distances are calculated with the equation:

$$d(\boldsymbol{V}, \boldsymbol{C}) = \sqrt{\sum_{i=1}^{n} (v_i - c_i)^2}$$

According to Mendez (2017, 50–51), Euclidean distance metric is not suitable for aggregating distances in high-dimensional setting, as the metric uses list-wise deletion to deal with missing values, which affects the recommendations. However, there are other ways to deal with missing candidate answers, such as assigning maximum distance between a candidate and a user for a question to which the candidate did not answer (this method does not prevent using Euclidean distance). Another problem with Euclidean and Manhattan distance metrics is that they both ignore possible correlations of VAA items by assuming them to be independent. Mahalanobis distance metric is suited to tackle this problem, as it takes into consideration these correlations (Katakis et al. 2014, 8). Mahalanobis can be used for calculating distance of two vectors (e.g., between a candidate and a voter), the equation for calculating the metric is

$$d(\boldsymbol{V}, \boldsymbol{C}) = \sqrt{(\boldsymbol{V} - \boldsymbol{C})^T S^{-1} (\boldsymbol{V} - \boldsymbol{C})}$$

where $\boldsymbol{V}$ and $\boldsymbol{C}$ are again voter and candidate vectors and S^{-1} is their inverse covariance matrix.

The inverse covariance matrix can be calculated by using all candidates' answers on VAA questions. Mahalanobis distance is not used as often as the other two distance metrics in calculating the VAA recommendations, even though it has provided more accurate recommendations than the other two methods (see Katakis et al. 2014). The reasons why this method has not been used more could be that firstly, it requires data (either candidate or user data) to form the covariance matrix which establishes the relations of issue variables. Secondly, average user cannot easily reconstruct their recommendation scores. The lack of training data for the covariance matrix is not an issue in the Finnish context, as there is a high number of candidate answers.

The final design choice concerns usage of weights. Wagner and Ruusuvirta (2012) point out that not all VAAs incorporate weights, but the majority of their sample of Western European VAAs incorporated weights. Weights can be either inbuilt to the VAA so that certain questions have a bigger role in the recommendation (design weights), or questions can be

weighted by parties/candidates (party weights), or alternatively weighting can be outsourced to users (user weights), who can choose the weights for individual questions according to their own preferences. The user weights are the most common of the three approaches. The idea behind the user weights is to allow users to incorporate their individual issue saliency into the recommendation. One way to implement this is to combine a 3-level variable (less important, neutral, more important) with each question, but there are also other (e.g., binary and 5-level) variants. Typically, the user weights are used by multiplying user–candidate distances for their respective questions. In practice, many users tend to ignore these weights (Wagner and Ruusuvirta 2012, 409). Romero Moreno et al. (2020, 6) also note that user weights can severely imbalance the VAA question composition, if there are some closely related issues in the VAA questions. The design weights, however, do specifically tackle this problem, for instance if there are two questions about the environment and three questions regarding immigration, these issues will have the same importance for the recommendation if the environment questions are multiplied with a factor of 1/2 and immigration questions with the factor of 1/3. It should be also noted that standard dimension reduction methods, such as EFA, do perform weighting of questions. However, the factorial solutions will be affected by issue-related imbalances of questions. The rarest of the three type of weights are the party weights which perform weighting of the questions based on parties' salience of each policy issue. Party weights can be assigned by experts or as Romero Moreno et al. (2020) suggest party manifestos could be utilized for this task.

2.4 The critique of voting advice applications

The critique of VAAs can be divided into two categories: methodological and normative. The methodological critique has been more apparent in the VAA research and it has pointed out that recommendations are usually dependent on multitude of design choices. Methodological critique of the VAAs seeks to point out that VAAs are not (yet) capable of providing uncontestable recommendations, as there is no single "right" design for a VAA that would provide perfectly objective and accurate recommendations. This means that recommendations provided by the VAAs are susceptible to change, if the design of the VAA or the set of questions are changed. Even though VAA recommendations might not ever be truly objective, the recommendations can be good approximations of the political landscape, and therefore be useful for choosing the right candidate. The degree to which VAA recommendations are

useful depends on how well the VAA in question counters the most apparent methodological concerns.

2.4.1 Methodological critique

Mendez (2017, 33) lists the most apparent methodological concerns of the VAAs:

1) statement formulation (Gemenis 2013; Van Camp et al. 2014)
2) the effects of statement selection (Lefevere and Walgrave 2014; Walgrave et al. 2009)
3) the configuration of answer categories (Rosema and Louwerse 2016; Baka et al. 2012)
4) the reliability of party coding (Gemenis 2013; Gemenis and van Ham 2014)
5) effects of algorithmic design choices, such as usage of different spatial logics (Louwerse and Rosema 2014; Mendez 2017)

I can point out at least three more issues, namely:

6) truthfulness of party positions (Wagner and Ruusuvirta 2012, 406; Gemenis 2013, 277)
7) handling missing candidate/party responses (Agathokleous et al. 2013)
8) scaling issues, which mean issues with forming ideological dimensions (Gemenis 2013)

In some VAA setups, some issues are mitigated whereas others are exacerbated. Next, I will discuss these methodological issues in more detail.

First, I am going to elaborate methodological issues related to statement formulation. Van Camp et al. (2014) have analyzed how VAA question-statements should apply survey methodology which embraces concreteness and avoid double-barrelledness, quantifications and qualifications. Concreteness refers to elimination of vagueness in the statements minimizing chances of misinterpretation. An example of a concrete statement would be that “Finland should shut down its coal-based energy production by 2025”, in contrast to “Finland should reduce its carbon emissions”. The first of the two provides a concrete issue statement with clear course of action, whereas the second one is not mentioning the specifics (how to reduce and in what timeframe). Gemenis (2013, 275) notes that too much specificity is also

not desirable. Double-barrelledness refers to statements that can be divided into multiple statements, this is to be avoided in formulating the VAA question-statements, as answers are no longer straight forward. Particularly troublesome are conditional statements where users might agree with the latter part of the statement, but disagree with the condition. However, Walgrave and Lefevere (2013) have pointed out that voters are susceptible to hold conflicting or unattainable preferences (e.g., oppose increasing taxes while wanting to increase government spending), in these situations it can be justified to bundle statements together (Gemenis 2013, 272). Quantifications and qualifications should also be avoided. Quantitative statements mean that something should be increased or decreased. Consider a statement: "military should be granted more funds". Disagreement with the statement has multiple interpretations, as disagreeing can mean preferring less funding or keeping the level of funding the same. Having quantitative statements measured by Likert scales contaminates the measurements, as more/less dimension is superimposed to response key agree/disagree (Gemenis 2013, 273–274). Qualifications are also problematic, as they frame statements by providing additional non-essential information. An example of a qualification in a statement would be the following: "The president should be in charge of foreign policy, including defining relations with Russia". Here, the qualification is the reference to Russia, which influences the answer of the respondent to be specifically related to that country. Therefore, qualifications should also be avoided. Following these statement formulation guidelines will lead to better VAA question-statements in terms of measurement validity and reliability.

The second and the most alarming methodological issue is about statement selection effects. Walgrave et al. (2009) note that statement selection is the central part of the VAA, as it is shown that different configurations of statements favor some parties over others, as also stated by the issue ownership theory (see e.g., Petrocik 1996). In their analysis of Belgian VAA data, Walgrave et al. (2009) ran 500 000 simulations by randomly selecting 36 statements out of 50 total statements and comparing which parties are most recommended to VAA users. Walgrave et al. (2009) noticed that some configurations of statements generated hardly any recommendations for certain parties, whereas other configurations increased recommendations for some parties. In a more recent study, Lefevere and Walgrave (2014) noticed that increasing the number of VAA questions related to cultural or economic dimensions increased the likelihood that parties with clear profiles on the dimensions got high recommendation scores. Center parties that did not hold extreme positions on these dimensions, rather mixed combination of policy stances, had a disadvantage (Lefevere and Walgrave 2014). Surprisingly, increasing the number of question-statements related to issues

that parties have issue ownership did not improve their recommendation scores (Lefevere and Walgrave 2014). The implications of these findings are that the dimensional structure of politics should not be overlooked while selecting the VAA statements. As it is evident that any set of VAA questions will have some inherent bias, the questions should relate to important issue dimensions determined by the electoral context, thus giving them theoretical foundation, which helps in differentiating parties from each other. Moreover, having multiple statements regarding each dimension will improve the validity of the voting advice (Lefevere and Walgrave 2014, 259). However, problems arise if VAA statement structure has not been planned and tested with the underlying dimensionality in mind. This might lead to situations where correlations emerge between multiple items unbeknownst to the VAA designers causing the calculated voting advice to be disproportionally influenced by positions on one ideological dimension, thus neglecting other dimensions of political conflict (see Kauppinen 2007, 141–142). An example of unbalanced dimensional structure is presented in figure 8.

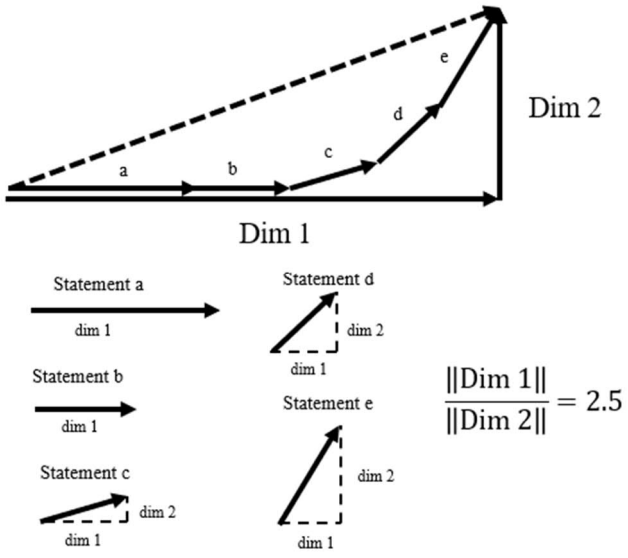

Figure 8. Unbalanced dimensional structure in VAA statements. Adopted from Kauppinen (2007).

Figure 8 presents a hypothetical case of a VAA with five statements that have an unbalanced structure. The five statements relate to two uncorrelated ideological dimensions (dim 1 and dim 2). Statements a and b relate to only one dimension (dim 1). However, statement b being

less indicative of the position on the dimension. Statements c, d and e are connected both dimensions. Statement c is primarily measuring position on dim 1, whereas statement d is equally connected to dim 1 and dim 2. Statement e is the only statement loading primarily on dim 2. After connecting all the statement vectors into a resultant vector (dashed line in the figure) one can break it into two components that are in line with dim 1 and dim 2. By comparing the lengths of these two vectors one notices that a resultant of statements that is in line with dim 1 is 2.5 times the length of a resultant in line with dim 2. This means that the overall composition of the VAA statements is focusing much more heavily on dim 1 than on dim 2, which is a structural imbalance resulting from statement selection. In some cases, this might be justified, if dim 1 is perceived to have more importance for the election. However, the emphasis on the dim 1 over dim 2 should be made apparent for the VAA users. It should be noted that actual VAA statements are much more complex, as typically more than two dimensions are surveyed in a VAA and the dimensions can also be correlated.

The third methodological issue concerns the configuration of answer categories. Rosema and Louwerse (2016) have shown that varying answering scale lengths do produce different recommendations. Rosema and Louwerse (2016) particularly tested for mechanical effects linearly transforming the answering scales. There can also be psychological effects which mean that respondents position themselves differently when the scale is changed, however this has not yet been studied in the context of VAAs (see Rosema and Louwerse 2016). Most commonly VAAs employ 5-point Likert scales, but there are also some cases with VAAs applying 2, 3, 11 and 101-point scales (Rousema and Louwerse 2016). Gemenis (2013) notes that Likert scales are particularly designed to be imperfect measurements of underlying latent variables, thus the questions need to be combined to extract respondents' positions on the latent variables, whereas 11-point scales (interval scales) measure respondents' positions on those dimensions directly. Despite its popularity, a Likert scale does not come without its own problems, Baka et al. (2012) have pointed out that VAA users had differing reasons to choose the middle-point answering option, namely lack of knowledge regarding the question or problematizing it.

The fourth methodological issue, namely the reliability of party coding is a specific issue related to VAAs that position parties based on expert evaluations. Expert evaluations are an effective way to prevent parties from manipulating their positions, but these evaluations are prone to errors. First, expert coders need to locate the party positions (e.g., from party manifesto) and second, transform the positions on VAA statements' 5-point response keys (Gemenis 2013, 277). Both steps of the process can contribute to coding unreliability, as

different document sources might lead to different positions from same parties and some questions might be too ambiguous leading to conflicting coding results (Gemenis 2013, 278). To improve the reliability of the party coding in VAAs, Gemenis (2015) has suggested using the Delphi expert survey method which is anonymous and iterative. Delphi method requires agreement between coders before positioning parties.

The fifth issue on the list concerns algorithmic design choices which have been introduced in detail in the previous chapter. Especially the effects of different spatial logics, number of policy dimensions and usage of different distance metrics have been researched (see Louwerse and Rosema 2014; Mendez 2017). There is no consensus on what elements the VAA algorithm should hold, however it is implied that Manhattan distance suits better to high-dimensional issue spaces, and Euclidian distance to low-dimensional issue spaces (Romero Moreno et al. 2020, 3). In candidate-based VAAs, there is much more flexibility in terms of algorithmic design choices in comparison to party-based VAAs, as candidates serve as free user-testers of the VAA statements. However, candidate-based VAAs (e.g., Finnish VAAs) have not been studied in the international VAA research and their full potential has not yet been uncovered.

The sixth issue concerns the truthfulness of candidates' VAA responses. The existence of such an issue has been identified (see e.g., Wagner and Ruusuvirta 2012, 406), but there is no data how prominent manipulative or strategic behavior is in candidate-based VAAs. Public scrutiny is expected to deter candidates not to fake their answers (Wagner and Ruusuvirta 2012, 406). Despite this, there is also evidence that at least one party (Feminist Party) has given guidance to their candidates on how they should answer to VAA questions in Finland (see HS 13.3.2019a; Aro 20.3.2019). One solution could be to use dynamic candidate profiles that augment the candidate VAA answers with social media sentiment data regarding the asked VAA questions (cf. Terán and Mancera 2017), although this solution is more vulnerable to outside manipulation, thus presenting an inside-outside manipulation trade-off.

The seventh issue is about missing candidate/party responses and how these should be handled when making the VAA recommendation. The issue of missing responses in VAAs has been researched by Agathokleous et al. (2013), the authors showed that recommendation accuracy is 50% even when 80% of the data is missing. Traditional party coding was among the most robust methods, whereas Mahalanobis classifier and decision trees performed poorly, as they are reliant on having accurate correlations between question items (Agathokleous et al. 2013). The authors do not provide any suggestions on how to replace missing values. Mendez (2017, 50–51) has suggested when using recommendation scores

ranging -100 to +100 one can simply divide the sum of matching weights with the candidate by the number of statements to which the user has responded. Alternatively, one can assign a maximal distance between a candidate and a user when candidate has not responded to the question. This method would penalize candidates for not answering, which should be communicated to the candidates before submitting their answers.

Scaling issues are the last major methodological issue. Scales, i.e. policy or ideological dimensions, are particularly important for VAAs that provide spatial maps. Typically, spatial maps are low-dimensional visualizations of policy positions that show both user's and candidates' positions (Germann et al. 2015). These dimensions can be constructed either deductively or inductively (Benoit and Laver 2012). Completely deductive approach defines dimensions *a priori* and then places political actors on the selected dimensions (Gemenis 2013, 280). Inductive approach selects issue-items and employs a scaling method (e.g., exploratory factor analysis) to assign items in dimensions (Gemenis 2013, 280). The problem with the inductive approach is that it is quasi-inductive, which means that it is reliant on statement selection, as constructed dimensions depend on the number and type of issues used in the scaling calculations (Benoit and Laver 2012, 207-208; Gemenis 2013, 280). However, Germann et al. (2015, 215) state that *a priori* positioning of items on scales without proper testing is the main reason for existing scaling issues. Scaling issues manifest via violations of unidimensionality requirement and lack of reliability on the constructed scale (Germann et al. 2015). Unidimensionality means that a single item should measure only a single latent trait. Germann et al. (2015) suggest that constructed dimensions should be validated dynamically with early user answers. In the context of candidate-based VAAs, early user data is not necessarily required, as candidates can be directly used for validation of the constructed scales, if the sample size is adequate (more than 2000 entries). Germann et al. (2015) suggest that Mokken scale analysis (Mokken 1971) and latent class reliability coefficient (van der Ark et al. 2011) should be applied for validating the VAA scales. Applying these methods can improve VAA recommendations in low dimensions and spatial maps that are shown to VAA users.

With smart design choices the significance of the methodological issues can be curtailed to a certain degree. The upshot from the methodological issues should be that VAAs are not objective "mirrors of politics" (see Fossen and van den Brink 2015), as the recommendations are always affected by subjective decisions made by the VAA designers to some degree. Fossen and van den Brink (2015, 350) note that even if all other issues would be solved, the effects of the statement selection on the recommendation are still too considerable and

systematic (see Lefevere and Walgrave 2014; Walgrave et al. 2009). Next, I will present the normative critique of the VAAs that contests inexplicit normative assumptions related to VAAs.

2.4.2 Normative critique

The normative critique of VAAs, firstly, points out the power imbalance between VAA users and developers (Fossen and van den Brink 2015) and secondly, contests the social choice model of democracy as the appropriate model for VAAs (Fossen and Anderson 2014). The power imbalance points to the fact that VAA developers are in charge of selecting the VAA questions, which is not a trivial task (see e.g., Walgrave et al. 2009). The developers can impose their own view of what is at stake in the election and voters can be misled by assuming that VAAs represent an objective view of politics (Fossen and van den Brink 2015). Fossen and van den Brink (2015) point out that VAAs are not mirrors, but they are dioramas providing one interpretation of the political reality. In particular, VAAs that provide low-dimensional spatial maps can be seen as political interventions, as they define the political landscape and present the user's position in the landscape (Fossen and van den Brink 2015, 352). One way to counter this imbalance of agenda-setting power could be to allow users themselves to define what is at stake in the election, in practice this would mean that users have more possibilities to tailor the VAA to match themselves to candidates in a political landscape of their own choosing (Fossen and van den Brink 2015, 353).

The second normative critique towards VAAs is regarding the presuppositions about democracy and citizenship that the voting recommendation is based upon. Fossen and Anderson (2014) note that the justification of a VAA does not merely rest on its technical capabilities, but it rests on these presuppositions. As VAAs provide voting advice there is an inbuilt assumption that the democratic process is primarily about elections (Fossen and Anderson 2014). Additionally, VAAs seek to improve citizens' competence to pick the "correct" candidate by improving their knowledge on candidate positions (Fossen and Anderson 2014). Fossen and Anderson (2014) name this type of a VAA design as the "matching VAA" which promotes the social choice model of democracy. Social choice theory perceives democracy as a function to aggregate preferences of individuals, which is widely contested by many theorists and citizens (Fossen and Anderson 2014). According to Fossen and Anderson (2014), there are also alternatives to this type of a VAA, as VAAs can be modified to serve different models of democracy (e.g., deliberative democracy) and perceptions of ideal citizen.

3 Research Objective and Questions

The research objective of this thesis is to envision a voting advice application design that would improve the quality of the voting advice. To reach this objective, I have listed three research questions. Two of which are prerequisite questions that pave way for the main research question. The first research question focuses on identification of problems, and the second one calls for a solution to one of the issues with the Finnish VAAs. Here are the prerequisite research questions:

> RQ1: What are the main problems with most popular Finnish VAAs?
> RQ2: How should VAA results be displayed to voters for them to gain insights regarding candidates' positions to other candidates?

After answering the prerequisite questions, I will address the main research question:

> RQ3: What kind of VAA design would solve identified issues with Finnish VAAs and provide voters more control over the VAA process?

4 Methodology

In this chapter, I will first present the overall methodological framework of the thesis. This is followed by a presentation of how the research data was collected. After which data analysis is discussed. The chapter is concluded by presenting the research data, namely five Finnish VAAs that were developed for the 2019 parliamentary election.

4.1 Research methodology and ethics

In terms of research strategy, this thesis applies design science research (DSR) approach (see van Aken 2015; Peffers et al. 2014). According to van Aken (2015, 1), DSR aims at "designing generic actionable knowledge to address types of field problems". Generic actionable knowledge refers to instrumental knowledge that can be applied to reach desired outcomes in the real world, and can be utilized in other contexts in comparison to where it was originally developed (van Aken 2015, 2). The objective of DSR is to improve the human condition by providing a design proposition as the main research output (van Aken 2015, 3). DSR is also set apart from other explanatory research strategies by its actor perspective, justification of the research based on pragmatic validity and focus on envisioning hypothetical scenarios. van Akken (2015, 1) states that DSR can be operationalized in many different research designs, including quantitative and qualitative designs (e.g., case studies, cross-case analyses). DSR-projects can vary in terms of their scope; however, a complete DSR-project has three parts: an explanatory part, a design part and a testing part (van Akken 2015, 1). In the first part, the field problem is analyzed and framed, in the second part, an intervention is designed and further developed, and in the final part, the suggested design is tested. In this thesis, first two parts of the design science research strategy are performed. Testing of the suggested design is left for future research.

As fitting for a DSR strategy, this thesis applies a case study design. The research data consists of Finnish VAAs at one point in time. The case study design allows for a holistic exploration of the phenomenon by using variety of data sources (Baxter and Jack 2008). Case study as a research design is distinguishable from other research designs by its focus on contemporary phenomena in its real-life context, especially when boundaries between phenomenon and context are not clear (Yin 1981). However, this thesis does not consist of only a single case, as there are five inspected VAAs, embedded in the same context (i.e., the context of 2019 Finnish parliamentary election). After introducing the five VAAs, a cross-

case analysis is performed. A cross-case analysis allows aggregation of results across cases and provides means to make generalizations based on the research data (Mathison 2005). Typically, this means that differences and similarities of cases are displayed visually (Mathison 2005), which is also performed in this thesis. By applying the case study design, I tested/used the VAAs and gathered VAA related publicly available information. For this research design to be properly implemented, it was imperative to have access to VAAs in order to evaluate them. It is also important to note that even though the thesis utilizes a case study design, it does not mean that the results derived from the Finnish case would not be applicable to other contexts. Thus, following the goal of design science research, my thesis produces generic actionable knowledge.

This thesis seeks to influence future development of VAAs, so it has a normative standpoint of creating an intervention to the current state of VAAs. Thus, the suggested VAA design envisioned in the thesis will not be value-neutral. Similarly, Flanagan et al. (2008) have pointed out that technological artefacts are not neutral, instead they embody values. My epistemological standpoint in this thesis is that no completely value-neutral and objective research exists, as researchers are embedded to the world that they study, and they can only perceive it through their own perspective. In a post-structuralist view, the perspective cannot be separated from the real world, as perspectives are integral parts of the world itself (see Devetak 2009). Therefore, it is the task of the researcher to make their own perspective and values apparent. My own perspective is one of a political scientist (being a doctoral candidate at the University of Helsinki). In respect to Finnish VAAs that are analyzed, I am an outsider that can inspect the VAAs from a standpoint of a typical user, as I have not been involved in development of any VAAs up to this date. As for values, my key values for VAA design are transparency and user empowerment, as I perceive that VAA users should be able to understand and influence the recommendation system that provides them voting advice, which can have real electoral consequences.

According to Flanagan et al. (2008), for technical design to incorporate values successfully the designers need to engage with technical, philosophical and empirical modes of inquiry. Technical mode refers to realization of design specifications, philosophical mode addresses the values to be reflected in the design, whereas empirical mode seeks to answer whether implemented values are embodied successfully, i.e. design achieves intended goals (Flanagan et al. 2008). From this perspective values should be made apparent, as they are driving forces of technological designs.

In terms of research ethics, no ethical boundaries were crossed in making this thesis. Even though individual candidate-level data was collected and used to answer one of the research questions, the main focus of this thesis were not candidates themselves. The candidate answers that were collected from Yle VAA were collected at nighttime, to minimize overusing Yle website servers at peak service hours. The created visualization that utilizes candidate data displays only candidates' names, party label and their matching scores to other candidates, thus not disclosing any sensitive information. All information that the created visualization utilizes is publicly available. Therefore, the thesis follows standard ethical guidelines for ethical research. Next, research data collection is presented.

4.2 Data collection

The research data for this thesis was collected by accessing five Finnish VAAs related to the 2019 parliamentary election. One and a half years after the 2019 election three out of five VAAs (Yle, Alma Media and MTV Uutiset) were still accessible and possible to use normally. Helsingin Sanomat VAA was no longer upkept, but candidates' VAA answers were accessible as they were attached to election results (see HS 2019). HBL still provided access to its VAA statements and showed party positions to the statements, but it no longer calculated voting advice (HBL 2019). To see how HBL VAA's voting advice looked like, I searched for screen capture images of its results, one of which revealed how the voting advice output looked like (see Sjöstedt 14.3.2020). With Helsingin Sanomat VAA I used VAA developer's video presentation that showed the UI elements in the results page (Futurice 10.5.2019).

An important part of the research data was the VAA statements. These statements were directly obtained from the VAAs. I also utilized news articles linked to the VAAs to gain additional information regarding the functioning of the VAAs that consisted of user guides, developer notes, etc. For this purpose, I searched VAA related content at the VAA platform providers' websites (using a keyword "vaalikone", meaning VAA). I obtained all articles mentioning VAAs few months prior to the election till election day. However, news articles were only taken in consideration, if they were related to the publisher's own VAA.

In order to answer the second research question, candidates' VAA answers were obtained from Yle VAA. Candidates' answers to VAA statements were collected from Yle VAA's candidate profiles with RSelenium R package (Harrison 2020) and these answers were cross-referenced with the Yle VAA dataset that was provided by the Yle editorial team per request.

Next, candidates' VAA answers were input back into the VAA to receive matching scores to other candidates. This was done because Yle had not disclosed their matching algorithm.

4.3 Data analysis

Following the DSR strategy, first explanatory part was executed by performing data analysis. To answer the first research question, five Finnish VAAs were analyzed by performing expert evaluation, i.e. qualitative content analysis, from a perspective of a typical VAA user, meaning that I utilized VAA data at hand, not contacting the VAA developers, but relying on publicly available information that all users can access. Harwood and Garry (2003) note that content analysis is a method for analyzing diverse data (e.g., textual and visual data). Content analysis is capable of comprising data into categories that allow for more convenient analysis and interpretation of the data (Harwood and Garry 2003). Qualitative content analysis is one form of content analysis (see Hsieh and Shannon 2005). Qualitative content analysis can be performed either deductively or inductively (Elo and Kyngäs 2008). In the deductive approach, the analysis is structured based on existing theories or observations, whereas in the inductive approach categories are derived directly from the research data (see Elo and Kyngäs 2008). In the analysis, I applied both inductive and deductive approaches.

The qualitative content analysis method of analysis was chosen over interviews with a selected group of layman users, which would not have necessarily provided any relevant suggestions on improvement of the VAAs, as most users lack domain knowledge on how VAAs operate and not knowing which parts of the VAA influence the voting advice output. The expert evaluation was carried out by testing the VAAs to identify user interface elements of the VAAs and test the VAA matching algorithms. Practical testing in combination with literature on VAA design was used to identify key issues with the Finnish VAAs to answer the first research question. Also, news articles were used for identifying possible issues (e.g., changes in the matching algorithm) that were not necessarily apparent by testing the VAAs.

Additionally, to answer the first research question, a mixture of quantitative and qualitative content analysis of VAA statements was performed to classify the statements based on which policy dimensions they were related to and count their frequencies. Qualitative content analysis was used to identify possible imbalances in the VAA statement structure by categorizing VAA statements. The analysis was performed in Microsoft Excel. Coding of the statements in respect to ideological dimensions and problems was performed based on predefined categories identified in the literature, i.e. deductively, whereas the statement types

were determined from the research data alone, i.e. inductively (see categorizations of statements in Appendix A). In the deductive approach, VAA statements were compared against ideological dimensions of Paloheimo (1988; 2005; 2008) and Suuronen et al. (2020). I coded the statements in two levels, one being the policy theme of the statements and the second one being the ideological dimension. Problems in the statements were identified based on Gemenis (2013) problem categories. In the inductive approach, I iteratively categorized the statements by their type (meaning that categories were revised as more statements were processed and statements were categorized again). After the categorization of the statements, frequencies of the categories were counted. The counts were then transformed into percentages by dividing the counts by the total number of VAA statements in each VAA. The focus on quantitative figures sets the analysis apart from typical qualitative analysis (see e.g., Morgan 1993). Therefore, I consider parts of the analysis to be a combination of qualitative and quantitative forms of content analysis.

4.4 Design development

Further following the DSR strategy, an effective intervention was designed. To answer the second research question, I took a role of a developer. After receiving candidate to candidate matching scores, I built a Shiny application (see Ylen vaalikoneverkosto 2019) in R programming language. The application allows users to explore closeness of candidates to each other by drawing a network based on candidate matching scores with visNetwork and shiny packages (see Almende et al. 2019; Chang et al. 2020).

To answer the third research question, I took a role of a designer and created paper prototypes. I utilized the answers provided to the first two research questions to envision a VAA design that embodies values of transparency and user empowerment, while mitigating the problems related to Finnish VAAs. The suggested VAA design also incorporates the developed network visualization method for candidate closeness.

4.5 Five Finnish VAAs

This thesis will focus on five Finnish VAAs: Yle (2019), Helsingin Sanomat (2019), Alma Media (2019), MTV Uutiset (2019) and HBL (2019). These VAAs were made accessible for the public approximately one month prior the election date (14 April 2019). Historically, these VAAs have been widely popular among the Finnish electorate. However, these are not

the only VAAs that were present during the 2019 election campaign period, e.g. Iltasanomat and Duunitori had their own VAAs. Iltasanomat VAA was co-created with Helsingin Sanomat VAA and it shared the same algorithm, but the questions were different (IS 18.3.2019). Yet, the five selected VAAs had the largest appeal and national audience, for example both Yle and Helsingin Sanomat VAAs attracted more than one million unique users during the campaign period (Yle 7.4.2019; HS 14.4.2019). The overall characteristics of the VAAs are represented in table 2. These characteristics are compiled based on my observations and VAA platforms' own source material regarding their VAAs. Next, I will introduce the VAAs.

Table 2. Five popular VAAs in the 2019 Finnish parliamentary election

	Yle	**Helsingin Sanomat**	**Alma Media**	**MTV Uutiset**	**HBL**
Number of candidate/party profiles	2265 candidates	2046 candidates	2149 candidates	circa 2000 candidates (?)	9 parties
Inclusion of all parties/candidates	Yes (all that answered)	Yes (all that answered)	Yes (all that answered)	Yes (all that answered)	No
Statements answered by	Candidates	Candidates	Candidates	Candidates	Parties
Number of statements	circa 30	30	25	35	25
Pro/against arguments	No	No	No	No	Yes
Response scales	5-point Likert scale*	5-point Likert scale	5-point Likert scale	0–100 (disagree/agree, slider)	4-point Likert scale + I don't know
Salience weights	No	No	Yes (binary)	No	Yes (select three statements)
Live match tracking	Yes	No	Yes	Yes	No
Voting advice	Candidate + Party	Party + Candidate	Candidate + Party	Candidate	Party
Visualizations	"Your parliament"	2-dimensional spatial map	2-dimensional spatial map	none	Bar plots
Algorithm	High dimensional (?) + Manhattan distance (?)	High dimensional + Manhattan distance	High dimensional (?) + ?	High dimensional (?) + ?	High dimensional (?) + ?
Pre-tested questions	Yes	Yes	No	No	No

*Note: * In Yle, candidates were not able to answer the middle point of the scale which is referred as "I don't know". However, the middle point works like a neutral answer and can be used by the users.*

Yle (Finnish Broadcasting Company) VAA was developed by Yle News Lab personnel in collaboration with researchers. Yle VAA had the highest coverage of candidate answers (90% of all candidates) of the five selected VAAs (Yle 14.3.2019b). Yle VAA had approximately 30 statements (depending on the district) 25 of which were national and five were tied to the district in question (Yle 14.3.2019b). All statements were presented in a single web page which appeared after the user chose their home municipality on the landing page. The user had an option to leave responses to statements empty.

Yle VAA was candidate-based with respect to inputs and output, as candidates provided their own answers to the statements and the VAA gave voting advice primarily regarding individual candidates. The main voting advice is presented as a list of candidates ranked in descending order with matching percentages (0–100%), while also displaying breakdown of seven issue-level matching percentages. The VAA output also included party recommendations on a second tab of results. Yle VAA asked users to place themselves on 5-point Likert scales, whereas candidates were only provided with a 4-point Likert scale without the middle-point option. The limitation of candidate answering options was explained by trying to prevent a "central answering strategy" where candidates purposefully place their answers in the middle of the response scale to increase their matching scores with all voters (Yle News Lab 14.3.2019). The VAA provided some short clarifications to some VAA question terms, but no pro/against arguments were provided to help users. Yle VAA did not provide an option to weigh the statements for users or candidates.

Yle VAA provided live match tracking on the upper banner of their web page which showed in real time 6–7 highest matching candidates with the user (Yle News Lab 14.3.2019). Yle VAA development team acknowledged that seeing the matches with candidates while answering the statements can influence the user, so users could click to hide the live match tracking, if they wanted to (Yle News Lab 14.3.2019). The usage of live match tracking was justified by noting that Yle VAA was designed to be "more of a research tool than a traditional online test" (Yle News Lab 14.3.2019). The VAA did not display any traditional visualizations, but it showed how parliament seats were divided among parties, if the user's closest matching candidates were chosen from each electoral district according to each district's number of seats. Yle did not disclose any information regarding their algorithm on their website, but according to my own experience with Yle VAA, the VAA utilizes high-dimensional matching with Manhattan distance, although this information has not been officially confirmed. Yle tested their VAA questions publicly before the launch and asked for citizens' feedback regarding the questions (see Yle 1.2.2019).

Helsingin Sanomat is the largest subscription newspaper of Finland (see KMT 2019). Helsingin Sanomat VAA was a candidate-based VAA, meaning that it collected answers from candidates directly. Helsingin Sanomat VAA had more than 2000 candidate profiles, the exact number of candidate profiles was obtained by collecting all accessible candidate profiles from Helsingin Sanomat VAA. The user experience in the VAA started with choosing the electoral district, after which VAA statements were revealed to the user one by one. The user had an option to skip a question, if they wanted. In total, Helsingin Sanomat VAA consisted of 30 statements, of which 10 were value statements and they have remained unchanged since 2012 (cf. HS 7.3.2019; HS 29.11.2017; HS 14.10.2012). Four out of the ten value statements were used to measure candidate positions on Left–Right dimension, another four on Liberal–Conservative dimension and the last two on Green–Non-green dimension (HS 14.10.2012; HS 7.3.2019). These statements were designed in collaboration with researchers Tuomas Ylä-Anttila and Teemu Kemppainen (HS 14.10.2012).

Helsingin Sanomat VAA did not provide any arguments for supporting user's decision-making processes while filling the VAA. The response scales for the Helsingin Sanomat VAA statements were on 5-point Likert scale. Salience weights were not available, nor was there any live match tracking to show to the user. Helsingin Sanomat VAA recommends a party as a primary voting advice for the user. This is a major point of difference to other Finnish candidate-based VAAs. However, Helsingin Sanomat VAA also provides matching to individual candidates regardless of the party, but this output is presented as a secondary piece of information in the results page. In the results page, user encounters first closest matching party and the best match within that party. The positional differences on five different policy topics are shown in detail accompanied with a question with the biggest difference between the user and the candidate. Additionally, Helsingin Sanomat VAA also visualizes user's and parties' positions on a two-dimensional spatial map that utilizes previously mentioned value statements. The VAA's matching algorithm is high-dimensional and calculates Manhattan distance between the user and a candidate. The algorithm was made publicly available (see HS 13.3.2019b), after which it was revised and explained (HS 18.3.2019). Helsingin Sanomat also pre-tested their VAA statements by placing them on public display and asked their readership for suggestions for new statements (HS 11.11.2018).

Alma Media is a Finnish media concern that owns many national newspapers (see Alma Media n.d.). Alma Media VAA (2019) contained answers from 2149 candidates (Uusi Suomi 1.4.2019). The VAA consisted of 25 questions of which five were tied to electoral districts (Uusi Suomi 1.4.2019). After choosing user location in the landing page the user was

presented with VAA question-statements one by one. The user had an option to skip a question by pressing the "skip question" button. Alma Media did not provide any additional information regarding the statements. The response scales to the statements were 5-point Likert scale with an option to use binary salience weights (more important/normal importance). Like the majority of the selected VAAs, Alma Media VAA had enabled live match tracking which allowed user to see maximally 20 highest scoring candidates ranked in a descending order while answering the VAA statements. What sets Alma Media VAA apart from other Finnish VAAs is the option to compare up to five candidates' answers to all VAA statements simultaneously to each other and to the user. In other Finnish candidate-based VAAs, only single candidate and user comparisons are permitted.

Alma Media VAA's voting advice was primarily regarding individual candidates and secondarily parties. This is evident in the order of the output elements, as candidate matches are shown first on the results page. On mobile devices, the candidate outputs are shown first, whereas on larger devices (personal computers, tablets, etc.) party and candidate results are shown side by side. The candidate focus of the VAA is also evident due to showing only three best matching parties without the ability to explore user's closeness to all parties. The VAA presented also two-dimensional spatial maps, where the user is placed on Left–Right and Liberal–Conservative dimensions with three closest and furthest candidates/parties. These maps are accompanied by an infographic that shows which questions were used to create the scales. The infographic also provided information on scale construction, which was carried out by averaging party positions district-wise, calculated through candidate answers. There is no explanation what algorithm Alma Media VAA uses to match users to candidates. However, my guess is that it uses high-dimensional matching algorithm. There is no source material which would indicate that Alma Media VAA had pre-tested their VAA questions publicly before launching the VAA.

MTV Uutiset is an online news media that is part of MTV media concern, which includes MTV3 television channel (MTV Uutiset n.d.). MTV Uutiset VAA gathered nearly 2000 individual candidate responses (MTV Uutiset 2019). This VAA is different from other selected VAAs, because it is embedded to a news article, instead of being on its own web page. MTV Uutiset VAA consists of 35 statements and a question regarding user's location (i.e., electoral district). The statements are presented to the user one by one. It should also be noted that users had an option to skip a statement. The response scales to VAA statements are on a 101-point Likert scale (i.e., 0–100 scale with end points at agree/disagree). The user operates a slider to pick their answer, although the user is not seeing the exact position (numeric value)

of the slider. MTV Uutiset VAA did not provide users with an option to apply salience weights.

MTV Uutiset VAA differed from other VAAs also in its simplistic design. There was no separate voting advice output section in the VAA, after inputting all the answers. The live match tracking simply freezes after answering all statements. The live match tracking showed users' matching percentages to all candidates in a descending order while responding to the VAA statements. There were no visualizations, nor were there any attempts to perform party-level matching. Also, the matching algorithm was left unexplained. My speculation is that the VAA utilized high-dimensional matching algorithm. MTV Uutiset had no published articles in which citizens could have pre-tested upcoming VAA questions or suggested any questions of their own.

HBL stands for Hufvudstadsbladet, which is the highest-circulation Swedish-language newspaper in Finland (see KMT 2019). HBL VAA was the only party-based VAA, as it asked nine major Finnish parties to answer the VAA statements. The party-focus was justified by pointing out that the votes always end up going to the party in the Finnish electoral system (HBL 2019). The party-focus was also endorsed by two political science professors, Åsa von Schoultz and Kimmo Grönlund, that collaborated with HBL on the project (HBL 2019). In HBL VAA, parties that did not have representatives in the parliament in the previous parliamentary term (2015–2019) were not included. In comparison to other VAAs, there was no question regarding user's location, as party answers apply to the whole country. There were 25 statements in total that were presented to the user one at a time. The statements were presented only in Swedish language. It is also noteworthy that the statements were selected from suggestions received from the readership (see HBL 2019; 22.1.2019).

Moreover, in contrast to other VAAs, user was also given short arguments for and against each statement, which were meant to help the user to answer. The answering scales were 4-point Likert scales (completely agree/disagree) with "I don't know" response alternative at the end of the scale. User could not skip a question. After answering all questions, user had an opportunity to choose up to three statements that they deemed important, which were then weighted. There was no typical live match tracking of parties, but the user could choose to reveal if any party picked the same response alternative, after the user had answered the question. However, this would prevent the user from altering that response without restarting the VAA. The output of the VAA exhibited matching scores (0–100%) of each party with the user, which were visualized as bar plots (see Sjöstedt 14.3.2020). No description of HBL

VAA's algorithm exists, yet I suspect that it uses high-dimensional matching. HBL did not publicly pre-test their VAA questions that they received from the readership.

5 Results

Results are presented in three parts. The first part responds to the first research question by identifying main problems with Finnish VAAs. The second part is devoted to developing and describing a new method for visualizing candidate closeness in VAAs, which answers the second research question. In the third and final part, I will answer the third research question by envisioning a new type of VAA design that utilizes answers to the previous research questions.

5.1 Identifying key issues with the Finnish VAAs

Five central issues, linked to VAA literature, were identified with 2019 Finnish VAAs. These issues are general in nature and concern all five inspected VAAs, although some issues are specific to candidate-based VAAs. The main issues are listed in order of importance:

1) Lack of transparency in VAA design
2) Lack of user interactivity with the VAA
3) Problems in VAA statement structure
4) Algorithmic issues
5) Lack of candidate comparisons and visualizations

Next, I will provide a thorough examination of each of these issues. I will also provide possible solutions to some of the issues.

5.1.1 Lack of transparency in VAA design

The first identified key issue with Finnish VAAs is lack of transparency. Transparency, openness, impartiality and methodological soundness are of central importance for VAA quality according to the Lausanne declaration (Garzia and Marschall 2014, 227–228). Over a decade ago, Kauppinen (2007) noted that output metrics provided by Finnish VAAs were lacking credibility, as users were not provided with an opportunity to get to know the functioning of the VAA or understand it. In addition, Kauppinen (2007) also noticed that Helsingin Sanomat was eager to explain the functioning of their VAA per request, whereas Yle and MTV Uutiset declined to provide information regarding their algorithms. In this respect, not a lot has changed, as I found out that the algorithmic transparency remains an issue with most

Finnish VAAs, also there are issues with design process documentation and handling of user and candidate data lacks transparency. My observations regarding these aspects are summarized in table 3.

Regarding algorithmic transparency, only Helsingin Sanomat had shared the code of the matching algorithm and explained its functioning after it received criticism (HS 10.3.2019; HS 13.3.2019b). Making the algorithm public revealed that the initial version of the algorithm was behaving unexpectedly, which sparked Helsingin Sanomat to revise their algorithm (HS 18.3.2019). All other VAAs had not publicly shared their algorithms, nor did they explain their functioning. This led to some confusion regarding their functioning, even for candidates. Jussi Saramo, a Left-Alliance candidate that got elected, criticized Yle for not properly explaining the effects of non-responses (Saramo 14.3.2019). By testing Yle (2019) VAA, it is evident that, if a candidate leaves a non-response to a question, not responding has a more negative impact on the matching score with the candidate than a complete disagreement (completely agree vs. completely disagree) between the user and the candidate.

Table 3. Transparency of five Finnish VAAs

		Yle	Helsingin Sanomat	Alma Media	MTV Uutiset	HBL
Algorithmic transparency	**Algorithm code available**	No	Yes	No	No	No
	Functioning of matching algorithm explained	No	Yes	No	No	No
Design process documentation	**Statement selection described**	Yes	Yes	No	No	Yes
	Stakeholders in creating the VAA are listed	Yes	No	No	No	Yes
User and candidate data handling	**User IP not linked to answers**	Not linked	Restricted access	No info	No info	Not linked
	Candidate answers shared publicly	Yes	No	No	No	Not applicable

In terms of documentation, three out of five VAAs described how statement selection was carried out (see HBL 2019; Yle News Lab 14.3.2019; HS 7.3.2019). Yle and HBL also described in detail which companies and individuals have been involved in creating the VAAs (Yle 14.3.2019a; HBL 2019). Handling of user and candidate data proved also to lack transparency. Yle (14.3.2019a) and HBL (2019) stated publicly that they did not link IP addresses or any identity revealing data of users to their answers. HBL (2019) also noted that user answers are not recorded. Other VAAs did not disclose what they do with users' answers or if they connect the answers to users' IP addresses or other identity markers. Helsingin Sanomat VAA was no longer active in 2020, which also prevented accessing their data policy page in the VAA (HS 2019). Alma Media (2019) noted in their VAA landing page that they will use users' location picked in the VAA, but not other user data, to personalize political advertisement for the users. In theory, user data could be used for political or consumer marketing purposes, therefore, it would be important to disclose how the data will be used and stored. Regarding the candidate answer data, only Yle was openly sharing candidate answers, even before the election date, although the candidate answers were anonymized for the public version of the dataset (Yle 5.4.2019). Researchers and journalists had a possibility to request access to the dataset without anonymization (Yle 5.4.2019). Sharing the candidate data would be important for users wanting to test the appropriate functioning of the matching algorithm, if this is also available. Also, sharing the candidate answers is important for democratic accountability of candidates, i.e. allowing candidates' stances on issues to be recorded and remain publicly accessible, even if the VAA itself would no longer be available.

In addition to aspects of VAA design transparency inspected in table 3, it could be possible to increase the credibility of a VAA's voting advice by sharing information about the extent in which parties (or parties of the best matching candidates) are recommended for the users in the VAA platform. It should be noted that this would need to be implemented with some precaution, as disclosing recommendation shares of parties based on user data has a potential for attracting malicious internet traffic on the VAA, as some actors might wish to discredit the VAA as biased or unreliable. Manipulation of the user recommendation metrics would simply mean overusing the VAA to make it seem that some parties have the most matches. Alternatively, users could inspect the credibility of the VAA in terms of how often candidates' own parties are recommended to them. This metric would not be susceptible for manipulation.

5.1.2 Lack of user interactivity with the VAA

The current versions of Finnish VAAs do not provide opportunities for users to interact and modify the VAA to their liking. VAA researchers have suggested more interactivity and responsibility for the users over a decade ago (see e.g., Kauppinen 2007). These calls for greater interactivity have been largely ignored by the VAA designers, as the 2019 Finnish VAAs provide only a single template of statements that the user can answer, also there are no options for users to choose between alternative response scales for individual questions or matching algorithms that determine how candidate–user distances are calculated. The uniformity of the VAA design for all users contradicts the reality that VAA users are a heterogenous group of voters. VAA users vary in terms of their political knowledge, interest towards politics, usage purpose and demographics (see e.g., van de Pol et al. 2014). Therefore, the idea of different users wanting different user experiences from the VAA is not unfounded. It could be that politically knowledgeable and interested users would be more interested to explore the VAA results in detail and tweak the VAA questions and algorithm according to their own liking. Users with less interest in politics and less domain knowledge could be more inclined to use the VAA with default settings and only quickly glance the VAA output. These are just examples, as there is no systematic survey data regarding VAA users' preferences on VAAs. However, facilitating custom-tailored VAA user experiences could result in wider public understanding that VAAs are always subjective constructs (see Fossen and van den Brink 2015) and this type of a VAA design would merely shift agenda setting power from VAA creators to users.

5.1.3 Problems in VAA statement structure

Here, I will inspect the statements of the five Finnish VAAs in terms of relevance to Finnish party competition. I will also categorize VAA statements by their type (political issue, personal value statement, etc.) and identify technical issues with the statements.

Haukio and Suojanen (2004) write as their first requirement for VAA statements that the statements should be relevant for the upcoming election. The relevance aspect can lure VAA developers to select only issues that are prevalent political issues discussed in the media. However, focusing on the day-to-day politics might overlook important aspects of political competition that are important for voters' vote choice. On the other hand, VAAs also influence the campaigning agenda which means that VAA developers have power to reinforce influence of already prevalent issues or give attention to issues neglected by the media.

My argument is that VAA statement composition should reflect underlying dimensions of political conflict of the country. VAA developers should not neglect these dimensions, but in contrast, they should try to keep dimensional structure balanced. Mere inclusion of user weights will not solve the structural imbalances in the VAA statements (see e.g. Kauppinen 2007). Moreover, Lefevere and Walgrave (2014) have pointed out that VAA statement selection is crucial, as any statement structures (emphasizing certain issues over others) favor some parties over others. Next, I will identify how statements of Finnish VAAs connect to relevant political dimensions by first introducing political science research on the topic and then categorizing VAA statements on the dimensions. My analysis does not however reveal, if the ideological dimensions are sufficiently measured by the statements. Isotalo et al. (2020) have found that 2019 Yle and Helsingin Sanomat VAAs did not encompass the whole range of the Left–Right dimension, as VAA statements did not make any differences apparent between communist and moderately leftist parties. This could be fixed by simply including more leftist VAA statements that bring out the differences of these parties.

According to Paloheimo (1988; 2005; 2008), there are seven ideological dimensions that are relevant for Finnish party competition: 1) Left–Right, 2) Center–Periphery, 3) National–International, 4) Elite–People, 5) Finnish speaking – Swedish speaking, 6) Conservative values – Liberal values and 7) Ecological values – Materialistic values. In Suuronen et al. (2020), anti-immigration and pro-immigration attitudes were identified to be a separate dimension. I combined both Paloheimo (1988; 2005; 2008) and Suuronen et al. (2020) dimensions, making the total number of dimensions eight.

I categorized VAA statements on the eight ideological dimensions based on the content of the statement (specifically based on the topic). It should be noted that thematic connection might not necessarily mean that there is a strong statistical connection. There were also some cases when a statement was linked to a dimension, but due to ambiguity of the statement, it was not always possible to determine which end of the ideological dimension was tied to agreeing or disagreeing with the statement.

In my categorization, (1.) Left–Right dimension was linked to the following topics: economy, social security, worker rights, privatization, outsourcing, Nato membership and basic income. The topics linked to the Left–Right were mostly socioeconomic. Kestilä-Kekkonen et al. (2018) have noted that Left–Right self-placement of voters is also correlated to topics beyond economic and class issues, such as EU, Nato membership and nuclear energy. However, I did not associate EU to Left–Right, as it is primarily International–National

dimension related, whereas nuclear power is linked to environmental values dimension. The full list of all categorized VAA statements is available in Appendix A.

(2.) Center–Periphery dimension measures attitudes on regional equality. This dimension was apparent in only one VAA statement, namely tertiary education availability in whole of Finland. Also, (3.) National–International dimension was rarely included in the VAA statements. The only topics associated with the dimension were about European union and development aid. (4.) Anti-immigration–Pro-immigration dimension was present in all VAAs. The dimension was linked to topics such as multiculturalism, refugees, citizenship, deportations, welfare chauvinism and foreign students. (5.) Elite–People dimension was not present in any of the VAA statements. This dimension would have captured populist attitudes, e.g., distrust toward elites and politicians. (6.) Monolingual–Bilingual dimension measures attitudes on role of Swedish language in Finland. The only question regarding the dimension was about reinstating Swedish as a mandatory subject in the matriculation examination. (7.) Conservative–Liberal dimension was surveyed in all VAAs, the dimension was linked to multiple libertarian issues that ask about increasing individuals' rights and minimizing the role of the state on everyday life (e.g., expanding alcohol selling rights, decriminalization of cannabis, right to euthanasia), law and order issues (e.g., police, security), rights of sexual minorities, gender equality and traditional values. (8.) Ecological–Materialist dimension measures attitudes on "green" issues, in the VAA statements such issues were meat consumption, flight tax, felling of forests, electric car subsidies, banning of regular cars and nuclear energy. Issues that were not linked to any of the mentioned dimensions were labelled as (9.) "Other". Other category included issues such as education, purchase of military equipment and hate speech. The share of statements belonging to each dimension for each Finnish VAA is represented in table 4.

Best represented ideological dimensions in the VAAs were Left–Right (28% on average across VAAs), Conservative–Liberal (22%) and Ecological–Materialist (19%). Immigration related dimension was also well represented, whereas National–International, Center–Periphery and Monolingual–Bilingual dimensions were almost non-existent in the VAA statements. Elite–People dimension was not represented in any of the five VAAs. On average 16% of the VAA statements were not linked to either of the eight dimensions. In relative terms, Alma Media and HBL VAAs were most successful in linking statements to relevant dimensions of political conflict by minimizing the share of statements in "Other" category.

Table 4. Shares of statements by ideological dimension, question type and problem type in five Finnish VAAs

	Yle	Helsingin Sanomat	Alma Media	MTV Uutiset	HBL
1. Left–Right	20%	27%	40%	23%	32%
2. Center–Periphery	0%	3%	0%	0%	0%
3. National–International	4%	3%	0%	0%	4%
4. Anti-immigration–Pro-immigration	8%	10%	10%	14%	12%
5. Elite–People	0%	0%	0%	0%	0%
6. Monolingual–Bilingual	0%	0%	0%	0%	4%
7. Conservative–Liberal	28%	17%	30%	20%	16%
8. Ecological–Materialist	16%	20%	15%	20%	24%
9. Other	24%	20%	5%	23%	8%
Total (n)	100% (25)	100% (30)	100% (20)	100% (35)	100% (25)
Political issue	88%	84%	85%	97%	100%
Personal value question	4%	10%	5%	0%	0%
Hypothetical question	0%	3%	0%	0%	0%
Opinion to non-issue	8%	3%	10%	3%	0%
Total (%) (n)	100% (25)	100% (30)	100% (20)	100% (35)	100% (25)
Ambiguous	24%	3%	10%	3%	0%
Double-barrelled state-ments	4%	7%	10%	3%	0%
Quantifications	8%	3%	5%	5%	12%
Qualifications	0%	0%	10%	0%	0%
Total (%) (n)	36% (9)	13% (4)	35% (7)	11% (4)	12% (3)

The worst performers in terms of diverse representation of political dimensions were Alma Media and MTV Uutiset VAAs which were lacking four of eight main dimensions. Helsingin Sanomat and HBL VAAs had the widest range of ideological dimensions covered, but neither of them had all dimensions suggested by Paloheimo (1988; 2005; 2008). In terms of issue

emphasis, Alma Media VAA was most heavily relying on only two dimensions Left–Right and Conservative–Liberal (70% of statements). In absolute terms, Alma Media VAA had similar number of statements in these two dimensions, but the relative share of those statements was amplified due to low number of total national questions.

By inspecting the type of the VAA statements, it is possible to detect that at least 84% of VAA statements in each VAA were about political issues (see table 4). HBL VAA was the most issue-centered of the five VAAs, as all HBL VAA's statements were related to political issues. It is desirable that the share of political issues is high, as VAAs are supposed to measure political positions (see Garzia and Marschall 2014, 227). Personal value questions and opinions to non-issues were also present in most VAAs. Personal value questions comprised of statements such as "I am willing to become a vegetarian to mitigate climate change" (Alma Media 2019) and "Traditional values are the basis of a good life" (Yle 2019). These types of questions survey respondent's values, not agreement with policy proposals. Opinion to non-issues were statements that failed to link to political issues or personal values (e.g., "A politician's main duty is to protect the interests of his or her constituents", Yle 2019). The last type of a VAA statement, a hypothetical question, was only present in Helsingin Sanomat VAA. This question asked whether regular cars should be banned, if self-driving cars would be safer (see HS 2019).

Lastly, VAA statements were screened for violations of standard statement formulation practices (see table 4). Gemenis (2013) has noted that VAA statements should be concrete, avoid double-barrelled statements and qualifications or quantifications. High share of ambiguous statements can erode the reliability of the VAA recommendations. This is so, because ambiguous statements have a tendency to hide ideological differences. Kestilä-Kekkonen et al. (2018) have noted that questions that are too generic and lack concreteness are not capable of revealing voters' actual Left–Right preferences, and secondly, because ambiguous statements can lead to misunderstandings or alternative interpretations between the user and the candidate. I categorized VAA statements as ambiguous, if they had competing alternative interpretations which would lead respondents to answer based on different criteria. Two examples of an ambiguous statement are "Traditional values are the basis of a good life" and "It is acceptable for some groups in society to be better off than others" (Yle 2019). These statements lack specificity, as they do not explain what the traditional values are, or which groups should be better off than others. One particularly ambiguous statement, which was present in nearly all VAAs, was regarding basic income. I labelled these statements to be ambiguous, if the statements did not specify the actual basic income model, as

it has been found out that Finnish citizens' support for basic income varies by basic income model ranging from 20 to 79 percent (see Pulkka 2020). Without clear specifications, basic income can be supported by the political left and right, even though their objectives associated with basic income are conflicting (scaling down welfare state vs. supplementing existing social benefit programs). From the five VAAs, Yle VAA had the most ambiguous statements (nearly a quarter of all statements in the VAA), which harms the overall reliability of Yle's voting advice. HBL did not have any ambiguous statements.

Double-barrelled statements, such as "Finland should downsize its tertiary education network and redirect funds toward top institutions and research" (Yle 2019), are combinations of multiple statements that can be broken down into individual statements. Alma Media VAA had in relative terms highest share of double-barrelled statements, whereas HBL did not have any.

Quantifications, such as "Finland should increase its refugee intake to at least 2000 people per year" (HBL 2019), are less/more statements that are asked to be evaluated by Likert scales. This is problematic, as disagreeing with the example statement can be interpreted as a preference to have less refugees or the same amount as currently. Quantifications were present in all VAAs, HBL having the most (12%) and Helsingin Sanomat the least (3%).

Qualifications, such as "Finland needs more nuclear power also after the construction of the fifth nuclear power plant, for instance to tackle climate change" (Alma Media 2019), seek to frame the question in terms of a specific condition. In the example, the desirability of nuclear power is boosted by framing it to be specifically used for fighting climate change. Statements with qualifications were present only in Alma Media VAA.

In terms of overall quality of the statements, Yle and Alma Media VAAs stand out. More than one third of the statements in both VAAs were identified to be problematic. Such a high number raises concerns regarding the reliability of these VAAs. Helsingin Sanomat, MTV Uutiset and HBL VAAs exhibited problems only in about 10% of the statements.

5.1.4 Algorithmic issues

First, the high-dimensional matching that is currently used in all five of the inspected Finnish VAAs disregards dependencies between questions. In practice, the dimensional structure of the data is not independent. If there are 30 questions in the VAA, the VAA matching algorithm treats the data as 30-dimensional (see e.g., Kauppinen 2007). Typically, this does not correspond with the reality, as multiple statements can belong to ideological and issue

dimensions, and within these dimensions the constituting statements correlate strongly with each other. If multiple highly correlated questions are asked in a single VAA and other questions show only weak correlations, this will impact the recommendations to be influenced heavily by some latent construct while other questions' contribution will be small in comparison. This issue could be tackled by either 1) performing dimension reduction and exposing the latent structure of the data by performing EFA, so that the issue space becomes low-dimensional or 2) calculating Mahalanobis distances that account for the correlational structure of the items.

Finnish VAAs have plenty of unused potential, as most European VAAs are party-based and they cannot test validity of constructed scales without user testing. In party-based VAAs, it is suggested that early user answers are analyzed and VAA is tweaked based on the results within one day after the launch (Germann et al. 2015), and this method has also been suggested for candidate-based VAAs (Kauppinen 2007). Uncertainty of construct validity is why high-dimensional VAAs have been so widely used, as VAA designers have been shying away from low-dimensional matching algorithms and spatial maps, as they seem too risky to implement with small number of data points (usually there are fewer parties than VAA questions). Finnish candidate-based VAAs do not have this issue, as there are over 2000 candidates in parliamentary elections, which is enough to validify scales. Finnish VAA designers could simply ask the candidates some extra questions to give themselves more possibilities for selecting the final questions that load according to expectations. However, performing EFA should not be an *ad hoc* solution, but the questions should be designed to be combined and not inspected in isolation. This is also the idea behind using the Likert scales (see Gemenis 2013). Having a predetermined dimensional framework as the basis of the item structure for the VAA would prevent imbalances resulting from statement selection. This, combined with EFA and calculating factor scores, would provide sound measurements for candidate and user positions from technical and political science perspectives.

Second, there have been concerns regarding the reliability of the party-level voting advice in the Finnish candidate-based VAAs. Helsingin Sanomat VAA made the headlines, as Feminist Party was recommended to many Green League and Left Alliance candidates, even though candidates' second closest matches were often their fellow party candidates (HS 10.3.2019). The original success of Feminist Party in the HS VAA was a result of having only a few candidates collectively deciding party's answers to the VAA questions (cf. Aro 20.3.2019) which resulted in nearly identical VAA answer profiles for Feminist Party candidates. This was a planned strategy to utilize "narrow" placement of candidates (alternative

strategy being a "wide" one, that maximizes the reach of the party). As party-level matching is given primacy in Helsingin Sanomat VAA, well-placed narrow placement of candidates maximizes user matches. In candidate-level recommendations, placing the candidates' answers widely is more advantageous. The criticism towards Helsingin Sanomat encouraged the VAA creators to change their party-matching algorithm by replacing party responses to be median responses of party's candidates instead of user's distances to all party candidates (HS 18.3.2019). The original party-matching algorithm did not work as intended, only 40 per cent of 2015 VAA candidates were recommended their own party, as higher number of candidates in a party was decreasing chances of bigger parties to be recommended (HS 18.3.2019). Using median of party's candidate responses increased the share of "correctly" recommended parties to candidates to 60 percent, whereas mean of candidate responses recommended 50 percent of the cases correctly (HS 18.3.2019). It should be noted that measuring VAA recommendation accuracy in terms of correctly assigning parties to candidates/users is a common method of evaluating VAAs (see e.g., Mendez 2017). However, there are issues with this metric, as firstly, parties can be ideologically heterogenous in terms of candidates and voter base, and secondly, as the number of parties increases so does the likelihood of overlapping in party positions. Therefore, 100% prediction accuracy is practically unattainable in multiparty systems. Next, I will present an overview of algorithmic possibilities and provide my own suggestion for calculating party-level matching scores with candidate-level data.

There are four main approaches in estimating party positions on VAA statements: 1) VAA creators can approach party headquarters and request official party answers directly, 2) use expert evaluations, 3) use party leader responses as proxies for party stances and 4) aggregate measures of party candidates. HBL used the first approach, whereas Helsingin Sanomat and Yle (also possibly Alma Media) used the aggregate measures approach. The aggregate measures approach can be implemented in two alternative ways. The first method is to calculate mean (or median/mode) of party's candidates' answers to each question, which can be represented mathematically $f(\frac{\sum_{i=1}^{n} x_i}{n})$, where f is the VAA algorithm and x_i is a candidate's answer and n is number of candidate responses by the party. The second method is to calculate mean (or median/mode) of candidates' recommendation scores for a party $\sum_{i=1}^{n} \frac{f(x_i)}{n}$. Party matching scores calculated with these two methods will most likely differ. Mean, median and mode are the most common measures of central tendency which seek to describe the underlying distribution. Each of these measures have their advantages and

disadvantages, the mean is influenced by the whole range of candidate responses, whereas the median is less affected by individual extreme answers, the mode on the other hand captures the most answered response which in one perspective could be considered to represent the party line. However, these aggregate measures have drawbacks. Mean and median values can indicate that a party has a neutral view on an issue, even if the party is split between two extreme views. Moreover, in some cases the majority of the party candidates' responses might not represent the official party line, which could be the case, if previously elected party representatives and current party leadership agree on an issue, but most party candidates that have not been elected or have little political power inside the party have a different stance on the issue. In this case, using any of the measures of central tendency would misrepresent the party's official stance on the issue.

My suggestion for determining the closest party for the user is to use multiple methods (if parties' direct responses are not available) and show all results to the user, even though this could complicate the interpretation of the results. One of the methods could utilize party leader closeness, one aggregate measures of candidates' answers by party and one aggregate measures of candidates' recommendations by party. I would also suggest including the following algorithm to the mix:

1) calculate individual candidates' matching scores to the user.
2) convert the matching score scale from 0–100% to -100–100, where neutral value is zero. The conversion of scores on the new scale is simply $2x - 100$, where x is the score in the original scale.
3) select the closest N candidates to the user, N meaning the maximum size of the party list in the district.
4) group closest N candidates by party and sum the matching scores.
5) select the party with the highest value.

The rationale behind this algorithm is that it does not claim to recreate party answers based on candidates' own answers to VAA statements, but it looks around the closest relevant candidates for the user. I view this approach as more realistic than trying to formulate party positions without party-level data. The only clear drawback of the algorithm is that it places parties with not full candidate lists at a disadvantage, but on the other hand, this reduces the effectiveness of small parties' coordinated placement of candidate answers close to each

other. I would not recommend using this algorithm as a stand-alone recommendation method for party-level matching, but it can complement other existing algorithms.

Kauppinen (2007) notes that there are two critical cases in VAA recommendations, when user is left without sufficient advice to choose between candidates. In the first case, a user has low matching scores with all candidates, and the second case is a scenario when a user matches with similar closeness with multiple candidates. To deal with the first case identified by Kauppinen (2007) the suggested algorithm converts the matching scores on the new scale. Without the new scale, summing candidate matching scores together would always increase the party sum, even though in reality candidates that are below 50% in the old scale are mostly disagreeing with the user on the issues. To respond to the second scenario of Kauppinen (2007), the algorithm looks beyond the closest candidate matches for the user, as it takes into account a larger number of relevant candidates. The number of relevant candidates can be altered, but the main idea in my formulation was to let this number vary according to the number of candidates that can be elected from the district (also known as district magnitude). On the other hand, not utilizing all candidates from all parties highlights the idea that not the whole party needs to be close to the candidate (e.g., there can be fringe candidates in the party or multiple opposing factions), but the important thing is that there is a sufficient number of candidates in a party that agree with the user. Alternatively, the user could be left to decide what they deem as a relevant number of close candidates. One more important aspect of the algorithm is that the closer the candidate to the user, the more it contributes to advocating the party to the user.

5.1.5 Lack of candidate comparisons and visualizations

None of the VAAs provide ways to compare individual candidates to each other in a comprehensive way. Alma Media (2019) provides an option to compare maximally five candidates side by side, which is a handy tool for comparing close matches with each other to find differences between candidates. However, there is no direct way to compare candidates' own matching scores with other candidates. Only visualizations that depict candidate positions with other candidates are in HS (2019) and Alma Media (2019) VAAs. They provide spatial maps, but these are centered around the user and do not show all candidates, but only a few closest and furthest ones. Outjes and Louwerse (2014) point out that any ideological dimensions that are presented should be tested to meet basic scalability requirements. There is no indication that any of the Finnish VAAs providing spatial maps would have tested their

visualized policy dimensions and that those dimensions meet scalability requirements. Additionally, these spatial maps utilize only a small fraction of actual questions related to ideological dimensions (see e.g., Alma Media 2019). To conclude, Finnish VAAs do not provide information on how candidates relate to each other in the VAA, thus the VAAs act as "black boxes" that reveal only parts of its data which is custom tailored to user inputs.

5.2 Visualizing high-dimensional VAA results with network graphs

Here, I answer the second research question: "How should VAA results be displayed to voters for them to gain insights regarding candidates' positions to other candidates?" by providing a novel network visualization. My suggestion is to use an interactive network graph to visualize candidate to candidate closeness by utilizing candidates' matching scores to other candidates. The suggested visualization solves the lack of transparency of the VAA results associated with VAAs that employ high-dimensional matching. When high-dimensional matching is used, users are only shown matching scores resulting from user–candidate distances, meaning that users are given only one view of the VAA results tied to their own perspective, while the rest of the VAA (candidate to candidate connections) remains hidden.

I built a web application by first collecting candidate to candidate matching scores from 2019 Yle VAA. The application is publicly available (see Ylen vaalikoneverkosto 2019). This application shows the user a network of each district's candidates matching scores. Candidates are represented as nodes in the network (size of the node is based on degree of the node). A tie is drawn between two candidates, if matching score between the candidates is equally large or higher than the user has specified. The minimum requirement for the tie to be drawn was set at 80% matching score level. The reason for such as high value was purely computational, as closely tied communities were noticed to take long time to settle, as the network utilizes physics simulations which allows nodes of the network to be moved around. The user can also choose, if they want to see candidates represented by the color of their party or community, which is calculated with a community detection algorithm. The user has also a possibility to select candidates by clicking on them directly from the network or by searching their name from a drop-down list. Figure 9 shows the overall view of the interactive network graph (see Appendix B for more illustrations of the web application).

The user can select the network to show only candidates of certain parties. In figure 10, only candidates from four parties are visualized. When a user selects a candidate, a list of candidate's five highest matches appears to the first tab of the draggable results box. The

second tab of the box presents how the candidate matches with other candidates by drawing distribution plots of the matching scores. These distributions along with the color of the ties in the network (lighter color indicates a higher matching score) reveal how well candidate matched with other candidates from their own or other parties. Figure 10 reveals that Katju Aro, Feminist Party leader, had a matching score of circa 90% on average with other candidates of the Feminist Party which is higher than with other parties. The exceptional closeness of Feminist Party candidates becomes apparent in the network graph by simply increasing the threshold of matching scores required for drawing ties between candidates. Thus, the interactive network graph reveals candidate to candidate closeness in a simple way. The web application also identifies possible strategic or manipulative behavior of parties by showing, if candidates of a certain party are providing identical answers to VAA statements (i.e., candidate to candidate matching scores are close to 100%).

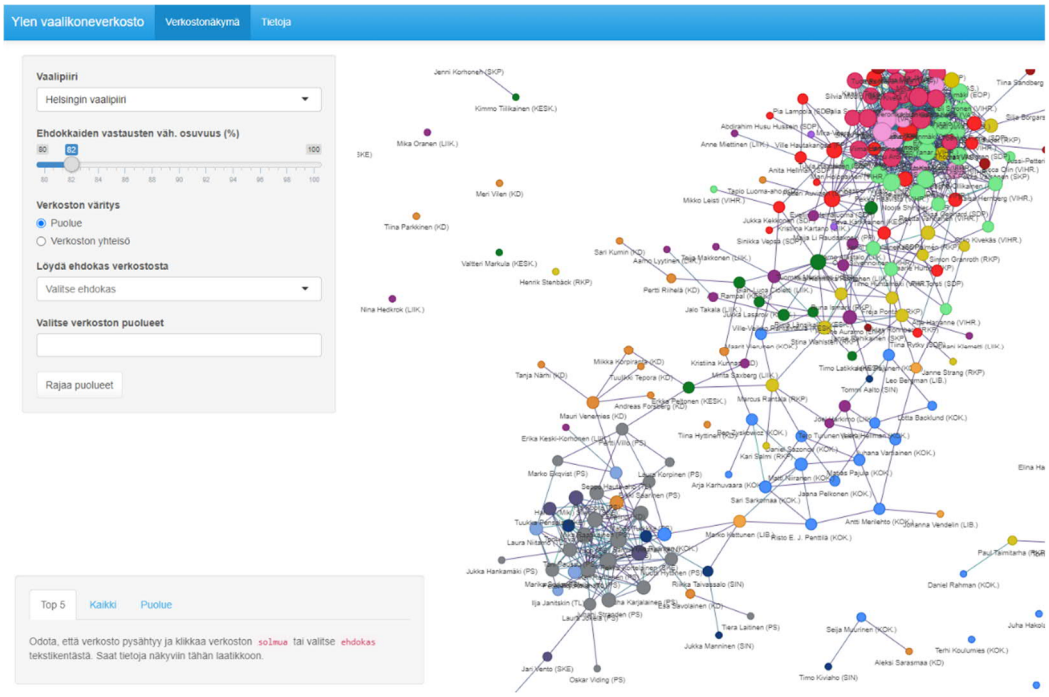

Figure 9. Overall view of the candidate-to-candidate network in Helsinki district. Screen capture from Ylen vaalikoneverkosto (2019).

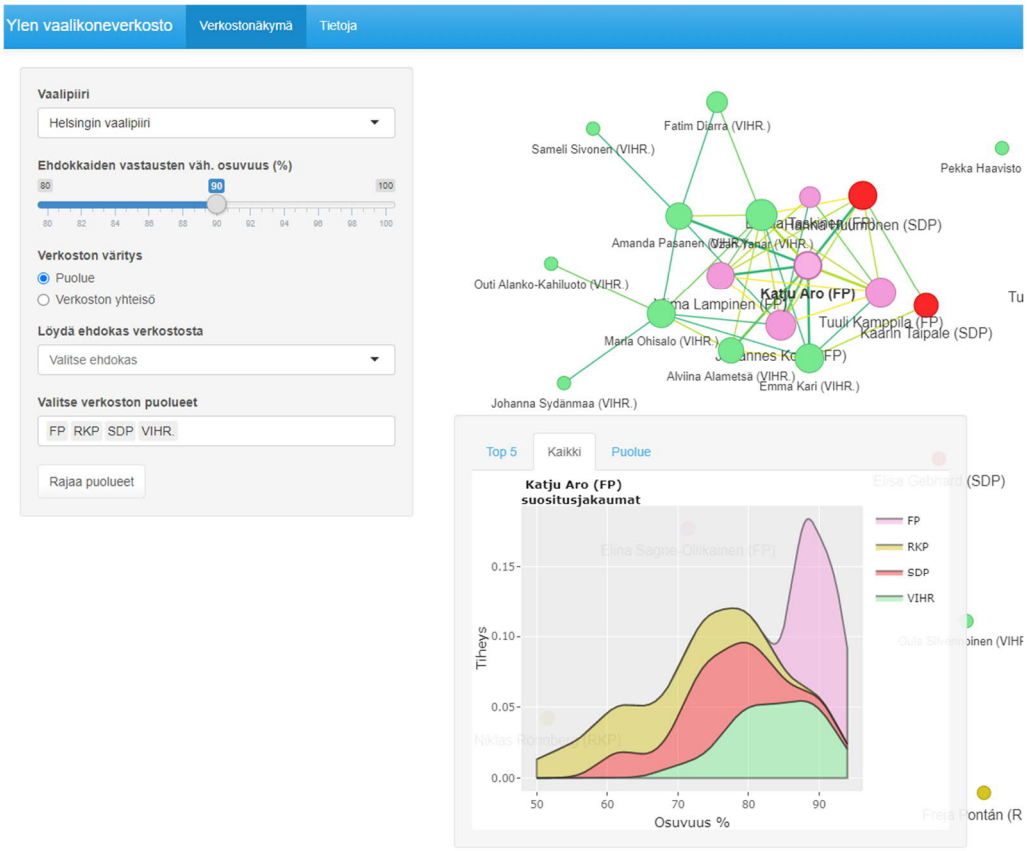

Figure 10. Candidate view of a pruned network with four parties, showing distribution of matching scores for candidate Katju Aro (FP). Screen capture from Ylen vaalikoneverkosto (2019).

5.3 Envisioning a new VAA design

In this part, I will answer the main research question, "What kind of VAA design would solve identified issues with Finnish VAAs and provide voters more control over the VAA process?", by proposing a two-step VAA design that increases user interactivity, design transparency and gives VAA users more influence on statement selection in national election VAAs. It should be noted that the envisioned design is most compatible with candidate-based VAAs covering national elections. However, parts of the envisioned design are also compatible with local and presidential elections.

5.3.1 First phase: Narrowing down the set of possible candidates

In the first step of my proposed VAA design, after users have picked their electoral district in the landing page, users have an option to identify and answer binary (yes/no) statements of crucial importance for them (i.e., "deal breakers"). The idea behind this optional first phase of the VAA is to filter out candidates that disagree with the user on the statements that

the user perceives important. The filtering follows a directional logic, as candidates need to be on the same side of the argument as the user. The rationale of narrowing down the set of possible candidates is to avoid a situation where the user being matched with a party/candidate with whom they have unreconcilable differences. The number of statements in this step should be high (for example, more than 50). To avoid overloading the voter by presenting all statements at once, the statements should be hidden by default, only showing voters different political themes under which individual statements locate. Users can use a search bar or click on the issue themes to view statements related to the theme, which makes the identification of the deal breaking statements an active process. The filtering of the candidates can also be visualized real time by indicating which candidates from which parties remain in the second phase of the VAA. The first phase design is illustrated in figure 11.

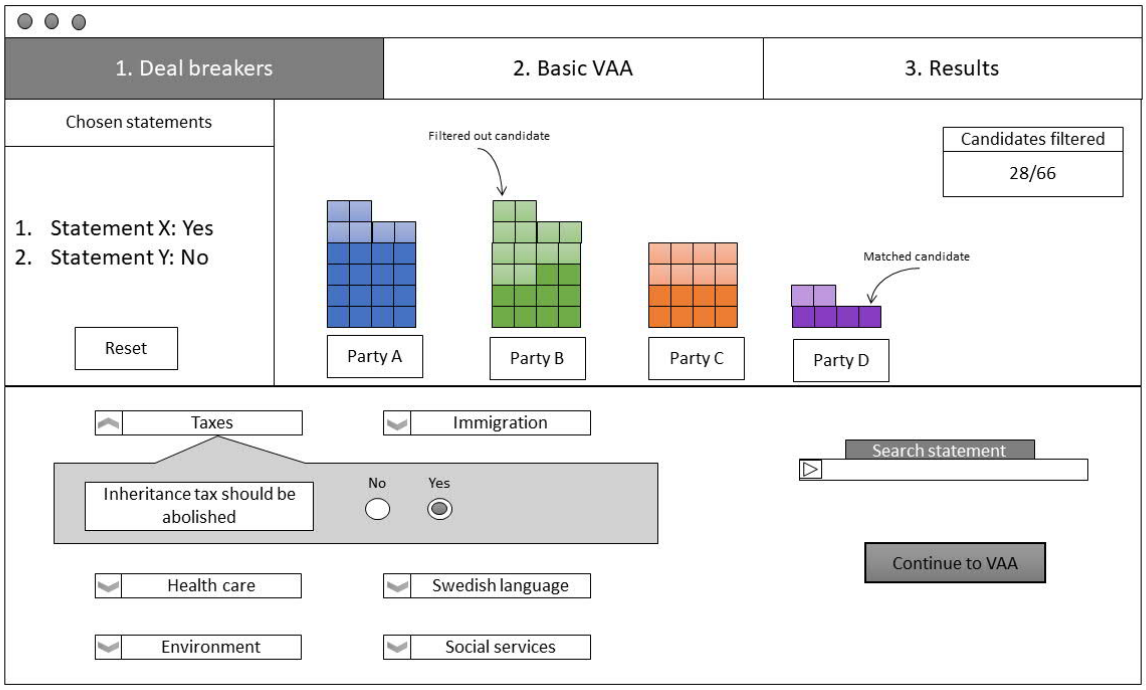

Figure 11. Illustration of the first phase in the envisioned VAA design that filters out disagreeing candidates.

To make the first step possible, candidates would need to be surveyed with a large number of yes/no statements. To limit the cognitive load of answering the VAA statements, candidates should not be required to provide comments to these statements. To encourage candidates to respond to the statements, it should be explained to them that abstaining from answering is treated similarly than disagreeing with the user, thus preventing access for the second step of the VAA for the user in question. As candidates cannot know which

statements individual users deem important, candidates should prefer to answer to all statements.

Instead of having VAA developers formulate all VAA statements, it would be preferable, if developers were to crowdsource this part of the statement selection. Having voters suggesting questions could increase influence of voters on VAA statements, which increases users' control over developers in the VAA process. Moreover, increasing voter influence on statement selection would decrease the existing power imbalances between VAA developers and VAA users, as voters would have a say in what is at stake in the election. Asking comments and suggestion for VAA statements has already been a standard practice for Yle, HBL and Helsingin Sanomat VAAs. Having the filtering phase in the VAA would enable more of the statements suggested by voters to find their way into the VAA.

5.3.2 Second phase: Performing a VAA for a smaller set of candidates

The second phase of the suggested VAA design consists of a VAA that matches a user with candidates/parties with approximately 30 statements that are all revealed to the user at the same time. Revealing the statements together allows users to compare, which statements they perceive more important than other statements. This has implications on users' ability to place salience weights.

In the suggested design, VAA statements are only related to (national) political issues to minimize ambiguity. The requirement of asking national-level issues is incompatible with local elections, as local elections would require statements to be related to local conditions which creates problems for keeping a balanced statement structure.

Also, it should be noted that live match tracking in this phase should be disabled not to bias users' answers. The design choice that sets my VAA apart from existing Finnish VAA designs is related to the matching algorithm. In my design, users are provided with algorithmic transparency and interactivity. For these algorithms to work, VAA statements are measured on Likert scales (e.g., 5-point or 11-point scales). It should also be noted that the algorithms are designed for candidate-based VAAs that perform matching based on candidates' direct answers to the VAA statements. The VAA user can choose in the beginning whether they want to get matches to parties or candidates. If the user selects party-level matching, then multiple operationalizations of party-matching scores will be calculated (see chapter 4.1.4), unless parties have input their official answers directly to the VAA.

Before answering the VAA statements, user picks the matching algorithm that they want to use. It is imperative that there is a short description of the algorithms in the VAA to guide the user. Also, links to the code of the algorithm and further documentation that explain how the algorithms work should be publicly available. The user can pick from two predefined algorithms: low-dimensional or high-dimensional, both of which are issue distance algorithms (see chapter 2.3.1). Also, two other algorithm types were considered (issue distance extended and learning VAA). However, both were disregarded, because including additional factors can be manipulated or systematically favor certain candidates. The problem with learning VAA is that it relies on the training process in which candidates'/users' answers are used to predict their party affiliation and this can be particularly challenging because the Finnish party system has more than ten parties and these parties are internally heterogenous in terms of ideological positions. In the chosen two algorithms users are linked to candidates solely based on VAA answers. Next, I will introduce these algorithms.

The default option is an algorithm that employs a low-dimensional modelling space and follows the proximity issue voting logic. In the default option, dimensions are constructed with applying exploratory factor analysis (EFA) on candidate responses. Using factor analysis has been proposed by van der Linden and Dufresne (2017), but they apply the method to pilot survey respondents that are voters. My method utilizes candidate responses to the VAA directly, as candidates' answers represent policy positions from which the voters can choose. Candidate and user positions on the dimensions are represented with factor scores. Distances between user and candidates are calculated with Euclidian distance metric. I chose low-dimensional matching to be the primary algorithmic option, as this type of matching does not overlook correlations between statements and it also provides material for visualizing results on spatial maps. Thus far, low-dimensional algorithms have not been employed in the Finnish VAAs, even though scale validation via Mokken scale analysis can be performed with candidate answers. Mokken scale analysis is a rather complex statistical procedure that can test whether items belong to constructed scales (see e.g., Sijtsma and van der Ark 2017). In the case of VAAs, Mokken scale analysis tests whether VAA statements belong to ideological dimensions. Low-dimensional algorithm would be best suited for a VAA that has a thought-out structure of statements which covers all ideological dimensions relevant to the political system. In the Finnish case, this would mean having at least three statements regarding all ideological dimensions suggested by Paloheimo (1988; 2005; 2008). This type of a VAA algorithm would make recommendations based on ideological positions, not simply based on aggregating differences in multiple issues. This means that users compare

their ideological positions to the candidates' positions. The output of the voting advice would be presented in terms of distances (not with regular matching scores), e.g. with bar-plots (or spider graphs). It should be noted that this algorithm is incompatible with presidential elections, as these elections have only a small number of candidates, thus presenting a problem for performing EFA.

The second predefined algorithm uses high-dimensional modeling space and Manhattan distance metric. This is currently the most popular method in matching users to candidates/parties (also in the Finnish VAAs). The user can choose whether it wants to use proximity, directional or hybrid distance matrices for evaluating differences in candidates' and user's responses. In the first algorithm, weighting of issues happens automatically as a part of construct the underlying ideological scales, but in the second algorithm no such weighting is applied. Therefore, it would be advisable to allow users to weight issues based on their issue salience. The weights should be placed next to each statement.

The second algorithm is best suited for VAAs that do not follow a planned statement structure that is tied to ideological dimensions. This type of an algorithm is also preferable, if user wants to emphasize some issues over others or does not want to answer most of the VAA statements. Non-responses of candidates and users are a bigger problem with the first algorithm, because for positions to be accurate on ideological dimensions, they require answers to all statements. In the high-dimensional approach, the effect of non-responses can be handled by setting the matching score scale to be from –100 to +100 and following the procedure suggested by Mendez (2017, 50–51).

An illustration of the suggested second phase VAA design is displayed in figure 12. In the example presented in figure 12, the user has chosen to match to candidates and utilize the high-dimensional algorithm that is called as "Issue-based matching" instead of the low-dimensional default option that is named as "Ideological position matching".

The voting advice output of the high-dimensional approach is typically reported as a list of candidates in descending order of the matching scores. To improve the user experience with the VAA output, I suggest that the list of candidates is placed in a table which presents basic information regarding the candidates (e.g., age, incumbency, home municipality) along with the matching scores. The table could be interactive, allowing the user to filter the table results based on selected criteria (e.g., candidate home municipality needs to match user's home municipality). However, improving the output list is not enough, as it still lacks a comprehensive view of the candidate positions. Moreover, the list does not reveal how close candidates are to other candidates. Lack of visualizations in high-dimensional

matching requires VAA users to laboriously compare individual candidates' responses side by side to see how these candidates differ. To solve this issue, I propose to use the developed interactive network visualization to reveal candidate ties to other candidates based on matching scores (see chapter 5.2).

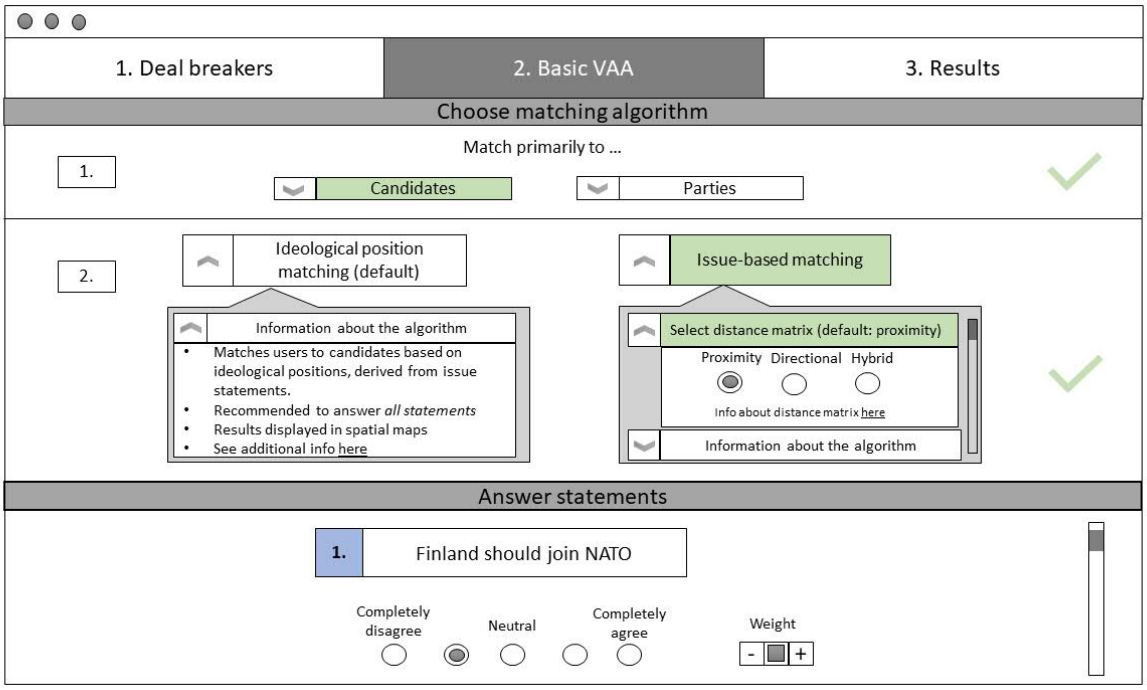

Figure 12. Illustration of the second phase in the envisioned VAA design that allows users to pick the matching algorithm.

6 Discussion

The discussion of this thesis starts by concluding the main results, after which implications of the research are discussed. This is followed by focusing on the validity of the results. Lastly, future research opportunities are presented.

6.1 Conclusion

In this thesis, it was shown that the five most prominent Finnish VAAs in 2019 parliamentary election were very similar in terms of response scales, algorithms and statements, even though there were some exceptions. The only major difference was whether voting advice was given primarily on parties or candidates. The VAAs analyzed in this thesis shared not only characteristics, but also major issues (RQ1), especially lack of transparency and interactivity. Also, shortcomings were detected in VAA statement structures and in the statement quality. Statement quality issues were most apparent in Yle and Alma Media VAAs. Algorithmic issues had been apparent in Helsingin Sanomat VAA. There was also a shortage of candidate to candidate comparisons and visualizations of the VAA results.

Problems regarding lack of transparency and interactivity have been plaguing Finnish VAAs for more than a decade (see Kauppinen 2007). Yet, no progress on developing the state of the VAAs has been made. In contrast, transparency of the VAAs has become more restricted as candidate answers are no longer shared openly with the public, which is a major disservice for democratic accountability. On a positive side, there were two promising signs in the 2019 VAAs that could lead to further improvement of the Finnish VAA landscape as a whole: 1) Helsingin Sanomat shared their matching algorithm with the public for the first time, 2) HBL VAA distanced itself from other Finnish VAAs by matching users directly with only parties. Regarding the first point, there is a chance that Helsingin Sanomat opening their algorithm for inspection might lead to other VAAs also sharing their algorithms in the following elections. HBL 2019 VAA was an important addition to Finnish VAAs, as party-based matching is the main VAA format internationally and it also provides important information on the actual parties' stances on issues, which could be applied for research purposes.

An interactive network visualization that presents closeness of candidates to other candidates (RQ2) was also envisioned and implemented. The visualization proved to be a useful tool for increasing transparency of the VAA results, especially if VAA in question employs

high-dimensional matching, as these types of VAAs typically hide inter-candidate matching information.

After identifying the state of the Finnish VAAs, I suggested a new VAA design that would increase VAA transparency, user interactivity and diminish power imbalance between VAA users and creators (RQ3). The suggested design matched users with candidates based on carefully selected statements that reflect the important ideological dimensions of political competition. The matching of users to candidates/parties was carried out in two phases. In the first phase, the user is given an option to filter out candidates based on statements that were suggested by voters prior to the launch of the VAA. In the second phase, user answers a set of questions that were chosen by the VAA developers. The user can choose from two different matching algorithms (low- or high-dimensional) depending on their preference, the default being a low-dimensional one. In the suggested design, having inbuilt default options is important as they prevent users from becoming overwhelmed by the complexity of the VAA design choices. Moreover, showing the complexities and the dependencies of VAA results on the algorithmic choices has a positive effect of pointing out that VAAs are inherently subjective tools. Instead of hiding the subjectivity aspect of the VAAs, it should be embraced by letting users personalize their VAA experiences in terms of selecting their own algorithms and statements.

6.2 Implications

Implications of this thesis are fourfold: 1) Finnish VAAs have issues, 2) these issues can be addressed with suggested VAA design, 3) the suggested design could provide richer data for political scientists about political candidates and 4) the suggested design has a clear potential for providing higher quality voting advice for VAA users, which in turn can help voters make more informed choices by easing candidate/party choice, thus having implications for democratic representation.

This thesis has shown that issues with Finnish VAAs are manifold and they seem to be persistent. After revealing the extent of problems with current Finnish VAAs, the question arises: can the voting advice of the Finnish VAAs be trusted to recommend the true closest candidates/parties for the user? Based on the findings of this thesis, my opinion would be no. This is also the view of the VAA developers themselves, as HBL and MTV Uutiset caution voters not to base voting decision solely on their VAA output (HBL 2019; MTV 2019). Helsingin Sanomat (7.3.2019) also notes that voters should explore multiple VAAs and not trust

a single VAA's results. Finnish VAAs can provide interesting results, but the lack of reliability, stemming from lack of transparency, unbalanced statement structures etc., erodes the usefulness of the recommendations.

In this respect, would the design suggested in this thesis be trustworthy enough? In my view, yes, as it tackles the most prominent issues with Finnish VAAs by providing transparency, user interactivity, balanced statement structure based on relevant ideological dimensions, improvements in algorithmic design and a new type of visualization. In terms of viability of the suggested design, Helsingin Sanomat, Yle and HBL already take in readers' suggestions for VAA statements, meaning that an infrastructure for receiving suggestions from the general public already exists. Formulating statements that reflect important ideological dimensions for Finnish party politics would require some expertise on the scale construction and consultation of political scientists, statisticians and journalists. Fortunately, Yle, HBL and Helsingin Sanomat have been collaborating with academics before, so I deem the overall design plausible.

For VAA design this thesis implies that VAAs should account for the local political context. Local political context consists of the electoral system that has implications for the selection of the matching algorithm, e.g. should the algorithm recommend parties or candidates. Local context is also reflected on political issues and ideological dimensions that are perceived relevant and should, thus be included in the VAA statements. Additionally, the thesis suggests that VAA design should be seen as a holistic enterprise, where all parts of the VAA should be seen important, starting form statement selection, to the user interface, matching algorithm and presentation of results. One should think how all these elements fit and interact together and not just select these elements of the VAA design in isolation.

Additionally, this thesis envisioned a novel way to match users to candidates based on closeness in ideological dimensions. Although the idea of using factor analysis to construct ideological dimensions to match candidates to users is not new, the originality of this approach was to construct ideological dimensions based on candidate answers without the need for pilot survey. This does not mean that in house testing of VAA statements should not be performed, but it implies that candidate-based VAAs could utilize candidates' answers to their full extent, leading into more reliable and valid recommendations by performing statement selection and weighting based on EFA.

Adopting the suggested design could have positive implications for political science, as VAA statements would better reflect the important ideological dimensions of party competition. This would provide more data for measuring candidate/party positions. Moreover,

including the first phase of the suggested design (filtering out candidates that have opposing opinions to deal-breaking questions) would mean more data on candidates' views, even if these statements would be recorded on binary answering scales. There is also a possibility to measure candidate responses to these statements in Likert-scales and transform the answers into binary scale later. This would allow for obtaining more fine-grained data about candidate positions on all statements.

Lastly, there are also implications for democracy. If the suggested VAA design were to improve the quality of the voting advice, this could mean that voters had a better opportunity to find their most preferable candidate. This could translate into better ideological and issue representation of voters' policy positions in the parliament, which is also the implicit goal of the matching VAA model. However, voters are not only interested in representation of issue positions, but they also care about demographical representation (see e.g., Pitkin 1967), meaning that candidates should fulfill certain criteria in terms of age, gender and locality. Moreover, including the possibility to filter candidate matches based on candidate characteristics in the VAA (e.g., candidate age, education level, political experience, number of personal votes in previous election) could allow voters to pick a candidate they truly want to vote for, while inspecting the VAA results. Currently, no such option exists, which means that voters need to process multiple cues regarding their vote choice separately. Vote advice from the VAA is from a voter's perspective only one piece of information that is considered in the voting decision-making process. In my suggestion, VAAs could be more than just providing one voting cue for the voters. VAAs have the potential to be platforms on which voters make their actual voting related decisions. To make this happen, VAAs should provide information regarding candidate characteristics and previous electoral performance along with ideological/issue closeness to the user with the opportunity to filter the pool of candidates, which could facilitate picking the ideologically closest candidate that matches voters' other criteria.

6.3 Validity of research

Validity of research can be evaluated in terms of internal and external validity. Internal validity reflects the confidence that can be placed on the obtained results. In terms of internal validity, it should be noted that having a single expert evaluator can be seen as a limitation. It could have been beneficial to survey multiple VAA experts' views on statement quality, UI and algorithmic design to increase the reliability of the findings. However, the evaluations

of the Finnish VAAs were linked to the existing literature, so one would expect the list of identified problems of multiple evaluators to be similar to the list presented in this thesis. Alternatively, one could have surveyed how typical VAA users interpret VAA statements. In this kind of a setup, statements with differing user interpretations could be identified and further explored. However, this was not the focus of this study.

One limitation with the suggested design is that allowing the public to suggest VAA statements and increasing interactivity of the VAA, do not inherently solve the issue of VAA developers having power over the statement selection process. However, the design can alleviate this issue by using the two-step process that utilizes more statements than regular VAAs, but within the matching VAA model one cannot fully escape the power imbalance of users and developers. Despite this issue, I still consider the suggested design to provide a major improvement to the current state of VAAs.

In terms of external validity, meaning the applicability of the results in other context, this thesis promoted some design choices that are only viable in national election candidate-based VAAs, however the overall design principles of increasing user interactivity and transparency along with some design elements were generalizable to all VAAs. The main limitation regarding external validity is lack of testing of the suggested VAA design and visualization.

6.4 Future research

For a future study, I propose creating the suggested VAA design and testing candidates' and users' opinions. Moreover, it would be interesting to study, if candidates suggested to the users change, depending on the inclusion of the first phase of the suggested VAA algorithm. This would reveal, if user empowerment is necessary in respect to statement selection, or can a typical VAA, consisting of VAA statements selected by the VAA developers, suggest fitting user–candidate matches. Moreover, voters could be surveyed about the type of the matching algorithm (low- or high-dimensional) that they prefer. Similarly, users' experiences could be recorded regarding the network visualization. The collection of user experiences does not have to be limited to survey data, but users' VAA usage could be observed in laboratory conditions to gain more fine-grained picture of users' behavior in filling the VAAs. These laboratory sessions could be complemented by having (group) interviews of users.

In addition, one could further develop the interactive network visualization by including additional variables regarding candidates (e.g., incumbency, home municipality, campaign

budget). These variables would affect the size or color of the network nodes. I would also suggest placing the user in the network based on their VAA answers. To maximize the user experience, the network graph should be in egocentric form, where the user is placed in the center of the graph.

The technological tools of today would allow for better VAA designs than the ones that are currently used. The goal of this thesis was to provide guidance to VAA designers and developers and encourage more innovative VAA designs. One should not limit possible VAA designs to the matching VAA model that takes citizens' positions on issues as given, but the future of VAAs should be embracing plurality of VAA designs. Fossen and Anderson (2014) suggest deliberative and agonistic VAA models that would either complement or even challenge matching VAA model. In the deliberative VAA model, the goal of the VAA would be to contest and update one's preferences on issues based on deliberative process. The agonistic VAA, instead aims to challenge the political landscape and problematize why certain issues are in the political agenda and some are not (see Fossen and Anderson 2014). The merit of the agonistic VAA model is that it recognizes the political nature of a VAA that automatically follows from the statement selection being in the hands of the VAA developers. Future research and VAA development should embrace these alternative models of VAAs and keep improving the matching VAA model designs. This course of action would strengthen democracy in multiple ways, easing candidate/party choice in elections, assisting voters in taking more informed stances on issues and make voters reflect what issues are important for them.

References

Agathokleous, M., Tsapatsoulis, N., & Katakis, I. (2013). On the Quantification of Missing Value Impact on Voting Advice Applications. In L. Iliadis, H. Papadopoulos, C. Jayne (Eds.), *Engineering Applications of Neural Networks*, *EANN 2013, Communications in Computer and Information Science, vol 383* (pp. 496-505). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-41013-0_51

Alma Media (2019). Eduskuntavaalit 2019: Vaalikone. Retrieved 7 October 2020 from https://www.uusisuomi.fi/vaalikone2019#/

Alma Media (n.d.). Alma Media lyhyesti. Retrieved 15 October 2020 from https://www.almamedia.fi/tietoa-meist%C3%A4/alma-lyhyesti

Almende, B. V., Thieurmel, B., & Robert, T. (2019). visNetwork: Network Visualization using 'vis.js' Library. R package version 2.0.9. https://CRAN.R-project.org/package=visNetwork

Anderson, J., & Fossen, T. (2014). Voting Advice Applications and Political Theory: Citizenship, Participation and Representation. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective* (pp. 217–226). Colchester, UK: ECPR Press.

Andreadis, I., & Wall, M. (2014). The impact of voting advice applications on vote choice. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective* (pp. 115–128). Colchester, UK: ECPR Press.

Aro, K. [@katjuaro]. (20.3.2019). Tweet. Retrieved 10 October 2020 from https://twitter.com/katjuaro/status/1108337411480829955

Baka, A., Figgou, L., & Triga, V. (2012). 'Neither Agree, Nor Disagree': A Critical Analysis of the Middle Answer Category in Voting Advice Applications. *International Journal of Electronic Governance, 5*(3-4), 244–263. https://doi.org/10.1504/IJEG.2012.051306

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. Qualitative Report, 13(4), 544–559.

Benoit, K., & Laver, M. (2012). The dimensionality of political space: Epistemological and methodological considerations. European Union Politics, 13(2), 194–218. https://doi.org/10.1177/1465116511434618

Bruinsma, B. (2020). Evaluating Visualisations in Voting Advice Applications. *Statistics, Politics and Policy, 11*(1), 1–21. https://doi.org/10.1515/spp-2019-0009

Budge, I., & Farlie, D. (1983). *Explaining and Predicting Elections: Issue Effects and Party Strategies in Twenty-three Democracies*. Allen & Unwin.

Campbell, A., Converse, P., Miller, W. E., & Stokes, D. E. (1960). *The American Voter*. New York: John Wiley.

Chang, W., Cheng, J., Allaire, JJ., Xie, Y., & McPherson, J. (2020). shiny: Web Application Framework for R. R package version 1.5.0. https://CRAN.R-project.org/package=shiny

Cho, S., & Endersby, J. W. (2003). Issues, the Spatial Theory of Voting, and British General Elections: A Comparison of Proximity and Directional Models. *Public Choice, 114*(3/4), 275–293. https://doi.org/10.1023/A:1022616323373

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1–9.

Devetak, R. (2009). Post-structuralism. In S. Burchill, A. Linklater, R. Devetak, J. Donnelly, T. Nardin, M. Paterson, C. Reus-Smit and J. True (Eds.), *Theories of International Relations* (4th ed., pp. 183–211). London, UK: Macmillan International Higher Education.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*(1), 20.

Djouvas, C., Antoniou, A., & Tsapatsoulis. N. (2018). *Improving Social Vote Recommendation in VAAs: The Effects of Political Profile Augmentation and Classification Method.* Paper presented at the 2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Zaragoza, Spain. https://doi.org/10.1109/SMAP.2018.8501885

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

Dumont, P., Kies, R., & Fivaz, J. (2014). Being a VAA candidate: Why do candidates use voting advice applications and what can we learn from it. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective* (pp. 145–160). Colchester, UK: ECPR Press.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Evans, J. A. (2004). *Voters and voting: An introduction*. SAGE Publications.

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. Oxford: Oxford University Press.

Flanagan, M., Howe, D. C., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In J. van den Hoven and J. Weckert (Eds.), *Information technology and moral philosophy* (pp. 322–353), Cambridge: Cambridge University Press.

Fossen, T., & Anderson, J. (2014). What's the point of voting advice applications? Competing perspectives on democracy and citizenship. *Electoral Studies, 36*, 244–251. https://doi.org/10.1016/j.electstud.2014.04.001

Fossen, T., & van den Brink, B. (2015). Electoral Dioramas: On the Problem of Representation in Voting Advice Applications. *Representations, 51*(3), 341–358. https://doi.org/10.1080/00344893.2015.1090473

Futurice (10.5.2019). Tech Weeklies: Creating a renewed voting advice application AKA "Vaalikone" (Video presentation). Retrieved 1 October 2020 from https://techweeklies.futurice.com/posts/creating-a-renewed-voting-advice-application-aka-vaalikone/

Garzia, D., & Marschall, S. (2014). The Lausanne Declaration on Voting Advice Applications. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective* (pp. 227–228). Colchester, UK: ECPR Press.

Garzia, D., Trechsel, A. H., & De Angelis, A. (2017). Voting Advice Applications and Electoral Participation: A Multi-Method Study. Political Communication, 34(3), 424–443. https://doi.org/10.1080/10584609.2016.1267053

Gemenis, K. (2013). Estimating parties' policy positions through voting advice applications: Some methodological considerations. *Acta Politica, 48*(3), 268–295. https://doi.org/10.1057/ap.2012.36

Gemenis, K. (2015). An iterative expert survey approach for estimating parties' policy positions. *Quality & Quantity, 49*(6), 2291–2306. https://doi.org/10.1007/s11135-014-0109-5

Gemenis, K., & van Ham, C. (2014). Comparing Methods for Estimating Parties' Positions in Voting Advice Applications. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective* (pp. 33–48). Colchester, UK: ECPR Press.

Gemenis, K., & Rosema, M. (2014). Voting Advice Applications and electoral turnout. *Electoral Studies, 36*, 281–289. https://doi.org/10.1016/j.electstud.2014.06.010

Germann, M., & Gemenis, K. (2019). Getting out the Vote with Voting Advice Applications. *Political Communication, 36*(1), 149–170. https://doi.org/10.1080/10584609.2018.1526237

Germann, M., Mendez, F., Wheatley, J., & Serdült, U. (2015). Spatial maps in voting advice applications: The case for dynamic scale validation. *Acta Politica, 50*(2), 214–238. https://doi.org/10.1057/ap.2014.3

Grofman, B. (1985). The Neglected Role of the Status Quo in Models of Issue Voting. The Journal of Politics, 47(1), 230–237. https://doi.org/10.2307/2131073

Harrison, J. (2020). RSelenium: R Bindings for 'Selenium WebDriver'. R package version 1.7.7. https://CRAN.R-project.org/package=RSelenium

Harwood, T. G., & Garry, T. (2003). An overview of content analysis. The marketing review, 3(4), 479–498. https://doi.org/10.1362/146934703771910080

Haukio, J., & Suojanen, M. (2004). Vaalikone poliittisena mediana. *Politiikka, 46*(2), 128–136.

Helsingin Sanomat (10.3.2019). Miksi HS:n vaalikone suosittaa yllättävää puoluetta? Vastauksia usein kysyttyihin kysymyksiin. Retrieved 10 October 2020 from https://www.hs.fi/politiikka/art-2000006030634.html

Helsingin Sanomat (11.11.2018). Lukijoiden vastaukset paljastivat terveyskysymyksemme huonoiksi – HS pyytää vielä apua vaalikonekysymysten testaamiseen. Retrieved 12 October 2020 from https://www.hs.fi/politiikka/art-2000005893417.html

Helsingin Sanomat (13.3.2019a). Feministit ovat aattellisesti hyvin lähellä vihreitä, mutta erojakin ehdokkaiden vaalikonekannoissa löytyy. Retrieved 13 October 2020 from https://www.hs.fi/politiikka/art-2000006033697.html

Helsingin Sanomat (13.3.2019b). HS:n vaalikoneen algoritmi avoimena: miten lasketaan samanmielisyys puolueeseen? Retrieved 11 October 2020 from https://www.hs.fi/politiikka/art-2000006033700.html

Helsingin Sanomat (14.10.2012). HS sijoitti kuntavaalien ehdokkaat arvokartalle mielipiteiden mukaan. Retrieved 13 October 2020 from https://www.hs.fi/politiikka/art-2000002571648.html

Helsingin Sanomat (14.4.2019). Yli 1,1 miljoonaa suomalaista on käynyt HS:n vaalikoneessa. Retrieved 13 October from https://www.hs.fi/politiikka/art-2000006070373.html

Helsingin Sanomat (18.3.2019). HS muutti vaalikoneen suositustapaa: uusi algoritmi avattuna ja selitettynä. Retrieved 10 October 2020 from https://www.hs.fi/politiikka/art-2000006039563.html

Helsingin Sanomat (2019). HS vaalikone: Eduskuntavaalit 2019. Retrieved 1 October 2020 from https://www.vaalikone.fi/eduskunta2019/

Helsingin Sanomat (29.11.2017). Vaalikone aiheuttaa ahdistusta poliittisen kentän oikealla laidalla – ehdokkaista ei tulosten perusteella löydy vapaamielistä oikeistolaista. Retrieved 13 October 2020 from https://www.hs.fi/politiikka/art-2000005469758.html

Helsingin Sanomat (7.3.2019). HS:n vaalikone aukesi: lähes kaikki puolueet ovat ottaneet askeleen vasemmalle. Retrieved 13 October from https://www.hs.fi/politiikka/art-2000006026006.html

Helsingin Sanomat (7.3.2019). Vaalikoneista on kaikkien syytä iloita. Retrieved 10 October 2020 from https://www.hs.fi/politiikka/art-2000006026015.html

Hosch-Dayican, B., Aarts, K., Amrit, C., & Dassen, A. (2013). Issue Salience and Issue Ownership Online and Offline: Comparing Twitter and Survey Data. *APSA 2013 Annual Meeting Paper, American Political Science Association 2013 Annual Meeting*. https://ssrn.com/abstract=2302137

Hotelling, H. (1990). Stability in competition. In A.C. Darnell (Ed.), *The Collected Economics Articles of Harold Hotelling* (pp. 50-63). New York, NY: Springer. https://doi.org/10.1007/978-1-4613-8905-7_4

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hufvudstadsbladet (2019). Valkompassen 2019. Retrieved 8 October 2020 from https://www.hbl.fi/valkompassen-2019/

Hufvudstadsbladet (22.1.2019). Delta i utformningen av HBL:s valkompass. Retrieved 14 October 2020 from https://www.hbl.fi/artikel/delta-i-utformningen-av-hbls-valkompass/

Inglehart, R., & Klingemann, H. D. (1976). Party identification, ideological preference and the left-right dimension among Western mass publics. In I. Budge, I. Crewe, D. Farlie (Eds.), *Party identification and beyond* (pp. 243-273). ECPR.

Iltasanomat (18.3.2019). IS muuttaa vaalikoneen suositustapaa. Retrieved 12 October 2020 from https://www.is.fi/politiikka/art-2000006039627.html

Isotalo, V., Mattila, M., & von Schoultz, Å. (2020). Ideological mavericks or party herd? The effect of candidates' ideological positions on intra-party success. Electoral Studies, 67. https://doi.org/10.1016/j.electstud.2020.102187

Iversen, T. (1994). Political Leadership and Representation in West European Democracies: A Test of Three Models of Voting. *American Journal of Political Science, 38*(1), 45–74. https://doi.org/10.2307/2111335

Jolliffe, I., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society of London. *Series A:*

Mathematical, Physical, and Engineering Sciences, 374(2065). https://doi.org/10.1098/rsta.2015.0202

Kamoen, N., Holleman, B., Krouwel, A., van de Pol, J., & de Vreese, C. (2015). The Effect of Voting Advice Applications on Political Knowledge and Vote Choice. *Irish Political Studies, 30*(4), 595–618. https://doi.org/10.1080/07907184.2015.1099096

Katakis, I., Tsapatsoulis, N., Mendez, F., Triga, V., & Djouvas, C. (2014). Social Voting Advice Applications-Definitions, Challenges, Datasets and Evaluation. *IEEE Transactions on Cybernetics, 44*(7), 1039–1052. https://doi.org/10.1109/TCYB.2013.2279019

Kauppinen, T. (2007). Vaalikoneiden tekninen toteutus ja kehittämistarpeet. In M. Suojanen and J. Talponen (Eds.), *Vallaton vaalikone* (pp. 127–156). SoPhi, 103.

Kestilä-Kekkonen, E., Tiihonen, A., & Westinen, J. (2018). Vääriä kysymyksiä vai vääriä vastauksia? Vasemmisto-oikeisto-ulottuvuus vuosien 2003-2015 vaalitutkimusten valossa. Politiikka, 60(2), 92–111.

KMT (2019). Levikkitilasto 2019. Retrieved 15 October 2020 from https://mediaauditfinland.fi/wp-content/uploads/2020/08/LT-tilasto-2019.pdf

König, P. D., & Nyhuis, D. (2020). Assessing the applicability of vote advice applications for estimating party positions. *Party Politics*, *26*(4), 448-458. https://doi.org/10.1177/1354068818790111

Kropko, J., & Banda, K. K. (2018). Issue Scales, Information Cues, and the Proximity and Directional Models of Voter Choice. *Political Research Quarterly, 71*(4), 772–787. https://doi.org/10.1177/1065912918760729

Lacy, D., & Paolino, P. (2010). Testing proximity versus directional voting using experiments. *Electoral Studies, 29*(3), 460–471. https://doi.org/10.1016/j.electstud.2010.04.008

Lau, R. R., & Redlawsk, D. P. (2006). How voters decide: Information processing in election campaigns. New York, NY: Cambridge University Press.

Lefevere, J., & Walgrave, S. (2014). A perfect match? The impact of statement selection on voting advice applications' ability to match voters and parties. *Electoral Studies, 36*, 252–262. https://doi.org/10.1016/j.electstud.2014.04.002

Lewis, J. B., & King, G. (2000). No Evidence on Directional vs. Proximity Voting. *Political Analysis, 8*(1), 21–33. https://doi.org/10.1093/oxfordjournals.pan.a029803

Lipset, S. M., & Rokkan, S. (1967). Cleavage Structures, Party Systems, and Voter Alignments: An Introduction. In S. M. Lipset and S. Rokkan (Eds.), *Party Systems and Voter Alignments: Cross National Perspectives* (pp. 1–64). New York: Free Press.

Louwerse, T., & Rosema, M. (2014). The design effects of voting advice applications: Comparing methods of calculating matches. *Acta Politica, 49*(3), 286–312. https://doi.org/10.1057/ap.2013.30

Macdonald, S. E., Rabinowitz, G., & Listhaug, O. (1998). On Attempting to Rehabilitate the Proximity Model: Sometimes the Patient Just Can't Be Helped. *The Journal of Politics, 60*(3), 653–690. https://doi.org/10.2307/2647643

Marschall, S. (2014). Profiling Users. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective* (pp. 93–104). Colchester, UK: ECPR Press.

Marschall, S., & Garzia, D. (2014). Voting advice applications in a comparative perspective: an introduction. In D. Garzia and S. Marschall (Eds.), *Matching voters with parties and candidates: Voting advice applications in comparative perspective* (pp. 1–10). Colchester, UK: ECPR Press.

Mathison, S. (2005). Encyclopedia of evaluation (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. https://doi.org/10.4135/9781412950558

Matthews, S. A. (1979). A Simple Direction Model of Electoral Competition. *Public Choice, 34*(2), 141–156. https://doi.org/10.1007/BF00129523

Mendez, F. (2017). Modeling proximity and directional decisional logic: What can we learn from applying statistical learning techniques to VAA-generated data? *Journal of Elections, Public Opinion and Parties, 27*(1), 31–55. https://doi.org/10.1080/17457289.2016.1269113

Merrill, S., & Grofman, B. (1999). *A Unified Theory of Voting: Directional and Proximity Spatial Models*. Cambridge: Cambridge University Press.

Mokken, R. J. (1971). *A Theory and Procedure of Scale Analysis: With Applications in Political Research* (Vol. 1). New York: De Gruyter, Inc.

Morgan, D. L. (1993). Qualitative Content Analysis: A Guide to Paths not Taken. *Qualitative Health Research, 3*(1), 112–121. https://doi.org/10.1177/104973239300300107

MTV Uutiset (2019). MTV Uutisten vaalikone julki – näin löydät itsellesi sopivimman ehdokkaan. Retrieved 3 October 2020 from https://www.mtvuutiset.fi/artikkeli/mtv-uutisten-vaalikone-julki-nain-loydat-itsellesi-sopivimman-ehdokkaan/7320524

MTV Uutiset (n.d.). MTV info. Retrieved 15 October 2020 from https://www.mtvuutiset.fi/yritys

Munzert, S., & Ramirez Ruiz, S. (2020). Meta-Analysis of the Effects of Voting Advice Applications. https://doi.org/10.31219/osf.io/utdn4

Munzert, S., Guess, A. M., Barberá, P., & Yang, J. (2020). Do Online Voter Guides Empower Citizens? Evidence from a Field Experiment with Digital Trace Data. https://doi.org/10.31219/osf.io/7prz4

Otjes, S., & Louwerse, T. (2014). Spatial models in voting advice applications. *Electoral Studies, 36*, 263–271. https://doi.org/10.1016/j.electstud.2014.04.004

Pajala, T., Korhonen, P., Malo, P., Sinha, A., Wallenius, J., & Dehnokhalaji, A. (2018). Accounting for political opinions, power, and influence: A Voting Advice Application. *European Journal of Operational Research, 266*(2), 702–715. https://doi.org/10.1016/j.ejor.2017.09.031

Paloheimo, H. (1988). Eduskuntavaalien 1987 kannatussiirtymät. *Politiikka*, 1/1988, 65–77.

Paloheimo, H. (2005). Puoluevalinnan tilannetekijät. In H. Paloheimo (Ed.), *Vaalit ja demokratia Suomessa* (pp. 202–228). Porvoo: WSOY.

Paloheimo, H. (2007). Vaalikoneiden käyttäjät vuoden 2003 eduskuntavaalikoneissa. In M. Suojanen and J. Talponen (Eds.), *Vallaton vaalikone* (pp. 57–76). SoPhi, 103.

Paloheimo, H. (2008). Ideologiat ja ristiriitaulottuvuudet. In H. Paloheimo and T. Raunio (Eds.), *Suomen puolueet ja puoluejärjestelmä* (pp. 27–60). Helsinki: WSOY.

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2014). A Design Science Research Methodology for Information Systems Research. *Journal of Management*

Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Petrocik, J. (1996). Issue Ownership in Presidential Elections, with a 1980 Case Study. *American Journal of Political Science, 40*(3), 825–850. https://doi.org/10.2307/2111797

Pianzola, J. (2014). Selection biases in Voting Advice Application research. *Electoral Studies, 36*, 272–280. https://doi.org/10.1016/j.electstud.2014.04.012

Pitkin, H. F. (1967). *The concept of representation*. Berkeley: University of California Press.

Pulkka, V. V. (2020). Perustulon kannatus Suomessa. *Yhteiskuntapolitiikka*, 85. http://urn.fi/URN:NBN:fi-fe2020050525340

Rabinowitz, G., & Macdonald, S. (1989). A Directional Theory of Issue Voting. *The American Political Science Review, 83*(1), 93–121. https://doi.org/10.2307/1956436

Romero Moreno, G., Padilla, J., & Chueca, E. (2020). Learning VAA: A new method for matching users to parties in voting advice applications. *Journal of Elections, Public Opinion and Parties*, 1–19. https://doi.org/10.1080/17457289.2020.1760282

Rosema, M., & Louwerse, T. (2016). Response Scales in Voting Advice Applications: Do Different Designs Produce Different Outcomes? *Policy and Internet, 8*(4), 431–456. https://doi.org/10.1002/poi3.139

Saramo, J. [@jussisaramo]. (14.3.2019). Tweet. Retrieved 12 October 2020 from https://twitter.com/jussisaramo/status/1106168061143461888?s=20

Shugart, M. S., Valdini, M. E., & Suominen, K. (2005). Looking for Locals: Voter Information Demands and Personal Vote-Earning Attributes of Legislators under Proportional Representation. *American Journal of Political Science, 49*(2), 437–449. https://doi.org/10.1111/j.0092-5853.2005.00133.x

Sigelman, L., & Buell, E. (2004). Avoidance or Engagement? Issue Convergence in U.S. Presidential Campaigns, 1960-2000. *American Journal of Political Science, 48*(4), 650–661. https://doi.org/10.1111/j.0092-5853.2004.00093.x

Sijtsma, K., & van der Ark, L. (2017). A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *British Journal of Mathematical & Statistical Psychology, 70*(1), 137–158. https://doi.org/10.1111/bmsp.12078

Singh, S. (2010). Contextual influences on the decision calculus: A cross-national examination of proximity voting. *Electoral Studies, 29*(3), 425–434. https://doi.org/10.1016/j.electstud.2010.03.014

Sjöstedt, J. [@jsjostedt]. (14.3.2019). Tweet. Retrieved 14 October 2020 from https://twitter.com/jsjostedt/status/1106294905784143881

Suojanen, M. (2007). Vaalikoneen lyhyt historia. In M. Suojanen and J. Talponen (Eds.), *Vallaton vaalikone* (pp. 13–28). SoPhi, 103.

Suomen virallinen tilasto (2019). Väestön Tieto- Ja Viestintätekniikan Käyttö [verkkojulkaisu]. Issn=2341-8699, Helsinki: Tilastokeskus. Retrieved 22 September 2020 from http://www.stat.fi/til/sutivi/2019/sutivi_2019_2019-11-07_kat_002_fi.html

Suuronen, A., Grönlund, K., & Sirén, R. (2020). Puolueiden äänestäjät. In S. Borg, E. Kestilä-Kekkonen, H. Wass (Eds.), *Politiikan ilmastonmuutos: Eduskuntavaalitutkimus 2019* (pp. 260–287). Oikeusministeriö.

Talponen, J., & Salminen, M. (2007). Vaalikone yksilöpäätöksenteon tukena yhteiskunnassa. In M. Suojanen and J. Talponen (Eds.), *Vallaton vaalikone* (pp. 39–56). SoPhi, 103.

Terán, L., & Mancera, J. (2017). Applying Dynamic Profiles on Voting Advice Applications. In Portmann E. (Ed.), *Wirtschaftsinformatik in Theorie und Praxis* (pp. 153–175). Wiesbaden: Springer Vieweg. https://doi.org/10.1007/978-3-658-17613-6_12

Thaler, D. (2015). *Issue voting with hearts and minds: Explaining heterogeneity in directional and proximity voting* (Doctoral dissertation). Available from ProQuest Dissertations Publishing.

Trechsel, A., & Mair, P. (2011). When Parties (Also) Position Themselves: An Introduction to the EU Profiler. Journal of Information Technology & Politics, 8(1), 1–20. https://doi.org/10.1080/19331681.2011.533533

Tromborg, M. (2019). Issue Salience and Candidate Position Taking in Parliamentary Parties. *Political Studies, 67*(2), 307–325. https://doi.org/10.1177/0032321718765520

Tsapatsoulis, N., Agathokleous, M., Djouvas, C., & Mendez, F. (2015). On the Design of Social Voting Recommendation Applications. *International Journal on Artificial Intelligence Tools, 24*(03). doi:10.1142/s0218213015500098

Uusi Suomi (1.4.2019). Alman vaalikone auki – Etsi ehdokkaasi Uuden Suomen sivuilla. Retrieved 14 October 2020 from https://www.uusisuomi.fi/uutiset/alman-vaalikone-auki-etsi-ehdokkaasi-uuden-suomen-sivuilla/933fc405-dba9-390c-88d2-360412ac340e

van Aken, J. (2015). What is Design Science Reasearch? an operationalization of DSR for the social domain in seven statements. Design Science Research institute.

Van Camp, K., Lefevere, J., & Walgrave, S. (2014). The Content and Formulation of Statements in Voting Advice Applications: A Comparative Analysis of 26 VAAs. In D. Garzia and S. Marschall (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective* (pp. 11–32). Essex: ECPR Press.

van de Pol, J., Holleman, B., Kamoen, N., Krouwel, A., & de Vreese, C. (2014). Beyond Young, Highly Educated Males: A Typology of VAA Users. *Journal of Information Technology & Politics, 11*(4), 397–411. https://doi.org/10.1080/19331681.2014.958794

van de Pol, J., Kamoen, N., Krouwel, A., de Vreese, C., & Holleman, B. (2019). Same but different: A typology of Voting Advice Application users in first- and second-order elections. *Acta Politica, 54*(2), 225–244. https://doi.org/10.1057/s41269-018-0083-3

van der Ark, L., van der Palm, D., & Sijtsma, K. (2011). A latent class approach to estimating test-score reliability. *Applied Psychological Measurement, 35*(5), 380–392. https://doi.org/10.1177/0146621610392911

van der Linden, C., & Dufresne, Y. (2017). The curse of dimensionality in Voting Advice Applications: reliability and validity in algorithm design. *Journal of Elections, Public Opinion and Parties, 27*(1), 9–30. https://doi.org/10.1080/17457289.2016.1268144

Vihtonen, E. (2007). Vaalikoneen synty ja kehitys yleisradiossa 1996–2006. In M. Suojanen & J. Talponen (Eds.), *Vallaton vaalikone* (pp. 29–38). SoPhi, 103.

von Schoultz, Å., & Papageorgiou, A. (2019). Policy or person? The electoral value of policy positions and personal attributes in the Finnish open-list system. *Party Politics*. https://doi.org/10.1177/1354068819891048

Wagner, M., & Ruusuvirta, O. (2012). Matching voters to parties: Voting advice applications and models of party choice. *Acta Politica, 47*(4), 400–422. https://doi.org/10.1057/ap.2011.29

Walgrave, S., & Lefevere, J. (2013). Ideology, Salience, and Complexity: Determinants of Policy Issue Incongruence between Voters and Parties. *Journal of Elections, Public Opinion and Parties, 23*(4), 456–483. https://doi.org/10.1080/17457289.2013.810630

Walgrave, S., Nuytemans, M., & Pepermans, K. (2009). Voting Aid Applications and the Effect of Statement Selection. *West European Politics, 32*(6), 1161–1180. https://doi.org/10.1080/01402380903230637

Westholm, A. (1997). Distance versus Direction: The Illusory Defeat of the Proximity Theory of Electoral Choice. *The American Political Science Review, 91*(4), 865–883. https://doi.org/10.2307/2952170

Yin, R. (1981). The Case Study Crisis: Some Answers. *Administrative Science Quarterly, 26*(1), 58–65. https://doi.org/10.2307/2392599

Yle (1.2.2019). Osallistu Ylen vaalikoneen tekemiseen: Testaa tässä eduskuntavaalien koneen väittämiä. Retrieved 13 October 2020 from https://yle.fi/uutiset/3-10623546

Yle (14.3.2019a). Toimittajalta: Vaalikoneessa ministeri ja lähikiskan Möttönen ovat samalla viivalla, ja siksi tuloksen pitääkin yllättää. Retrieved 11 October 2020 from https://yle.fi/uutiset/3-10683701

Yle (14.3.2019b). Etsi ehdokkaasi Ylen vaalikoneessa – mukana noin 2 200 ehdokasta koko maasta. Retrieved 13 October 2020 from https://yle.fi/uutiset/3-10688075

Yle (2019). Welcome to Yle's election compass! Retrieved 1 October 2020 from https://vaalikone.yle.fi/eduskuntavaali2019?lang=en

Yle (5.4.2019). Nyt tutkimaan: Ylen eduskuntavaalien vaalikoneen aineisto julkaistu avoimena datana. Retrieved 11 October 2020 from https://yle.fi/uutiset/3-10725384

Yle (7.4.2019). Ylen vaalikonetta käytetty yli miljoona kertaa – oletko sinä jo löytänyt ehdokkaasi? Retrieved 13 October 2020 from https://yle.fi/uutiset/3-10726545

Yle News Lab (14.3.2019). Näin keksimme vaalikoneen uudestaan – 5 keinoa tehdä ehdokkaan valitsemisesta kiinnostavampaa. Retrieved 11 October 2020 from https://newslab.yle.fi/blog/3lpCALgzagtq3gxVchdkNv

Ylen vaalikoneverkosto (2019). Ylen vaalikoneverkosto. Retrieved 15 October 2020 from https://veikkoisotalo.shinyapps.io/vaalikoneverkosto-app/

Ylä-Anttila, T. (2014). Perussuomalaisten sisäiset poliittiset suuntaukset: Julkisen oikeuttamisen analyysi. *Politiikka, 56*(3), 191-209.

A. **VAA statements**

Table A1. Alma Media (2019) VAA statements

Problem	Type	Ideological dimension	Topic	Statement
-	Political issue	<< Left–Right >>	Economy	1. Oikeutta poliittisiin lakkoihin pitää rajoittaa.
-	Political issue	<< Left–Right >>	Equality (economic)	2. Toisen asteen koulutuksesta oppimateriaaleineen (lukiot ja ammatillinen koulutus) pitää tehdä täysin maksutonta.
-	Political issue	<< Left–Right >>	Economy	3. Pääomatulojen verotusta pitää kiristää.
-	Political issue	<< Conservative–Liberal >>	Libertarianism	4. Viinejä pitää saada myydä vapaasti ruokakaupassa.
-	Political issue	<< Ecological–Materialist >>	Traffic emissions	5. Suomen pitää seurata muiden Pohjoismaiden esimerkkiä ja asettaa aikaraja uusien polttomoottoriautojen kieltämiseksi.
-	Political issue	<< Left–Right >>	Economy	6. Perintöverosta pitää luopua.
Double-barrelled	Political issue	<< Left–Right >>	Elderly care & taxes	7. Vanhustenhoidon laatua pitää nostaa tuloverotusta kiristämällä.
Double-barrelled	Political issue	<< Left–Right >>	Economy & social security	8. Suomen valtion velkaantumista pitää vähentää leikkaamalla sosiaalitukia.
Qualification	Political issue	<< Ecological–Materialist >>	Nuclear energy & climate change	9. Suomi tarvitsee lisää ydinvoimaa myös viidennen ydinvoimalan jälkeen muun muassa ilmastonmuutoksen torjuntaan.
-	Political issue	<< Immigration >>	Immigration	10. Suomen pitää pyrkiä rajoittamaan jyrkästi turvapaikanhakijoiden pääsyä maahan.
-	Political issue	<< Conservative–Liberal >>	Equality (gender)	11. Vanhempainvapaa pitää jakaa tasan naisten ja miesten kesken.
Ambiguous	Political issue	<< Left–Right >>	Social security (basic income)	12. Kansalaisen pitää saada tulevaisuudessa perustuloa ilman ehtoja
-	Political issue	<< Left–Right >>	Security policy	13. Suomen pitää rajoittaa sotilaallista harjoittelua Naton ja Yhdysvaltojen kanssa.
Quantification	Political issue	Other	Defence policy	14. Hornetien korvaaminen on niin kallista, että uusia hävittäjiä voidaan ostaa vähemmän kuin puolustusvoimat haluaa.
Qualification	Political issue	<< Conservative–Liberal >>	Equality (gender)	15. Kotihoidontuen maksimiaika on pudotettava kolmesta vuodesta kahteen, jotta naisia saadaan nopeammin takaisin työmarkkinoille.
-	Political issue	<< Conservative–Liberal >>	Sexual minorities	16. Suomessa pitää hyväksyä virallisesti kolmas sukupuoli.
-	Political issue	<< Conservative–Liberal >>	Church and state	17. Evankelisluterilaisen kirkon verotusoikeus pitää poistaa.

-	Opinion to non-issue	<< Conservative–Liberal >>	Equality (gender)	18. Sukupuolten välinen palkkatasa-arvo on toteutunut hyvin suomalaisessa työelämässä.
Ambiguous	Opinion to non-issue	<< Immigration >>	Immigration	19. Maahanmuuttajan sopeutuminen suomalaiseen yhteiskuntaan riippuu kulttuuritaustasta ja uskonnosta.
-	Personal value question	<< Ecological–Materialist >>	Vegetarianism & climate change	20. Olen valmis ryhtymään kasvissyöjäksi ilmastonmuutoksen hidastamiseksi.

Table A2. HBL (2019) VAA statements

Problem	Type	Ideological dimension	Topic	Statement
-	Political issue	<< Ecological–Materialist >>	Flight tax	1. Finland ska införa en flygskatt
-	Political issue	<< Ecological–Materialist >>	Meat consumption	2. Politisk styrning ska få oss att äta mindre kött
-	Political issue	<< Ecological–Materialist >>	Traffic emissions	3. Försäljningen av nya bensin- och dieselbilar ska förbjudas senast på 2030-talet
-	Political issue	<< Ecological–Materialist >>	Recycling	4. Det behövs en lag om att sortera plast i hela landet
-	Political issue	<< Ecological–Materialist >>	Forest ownership	5. Skogsägare som ökar kolsänkan ska få ersättning från staten
-	Political issue	<< Left–Right >>	Equality (economic)	6. Böcker och annat material i gymnasier och yrkesskolor ska vara gratis för alla studerande
-	Political issue	Other	Education	7. Den nya läroplanen lägger för mycket ansvar på eleven
Quantification	Political issue	Other	Education	8. Elever vid yrkesskolor ska ha rätt till minst 25 timmar lärarledd undervisning i veckan
-	Political issue	<< Left–Right >>	Social benefits	9. Studiestödet ska höjas
-	Political issue	<< Left–Right >>	Inheritance tax	10. Arvsskatten ska slopas
-	Political issue	<< Left–Right >>	Taxation of capital	11. Kapital ska beskattas mer än nu
-	Political issue	<< Left–Right >>	Unemployment benefits	12. Det ska krävas en motprestation av den arbetslösa för att få dagpenning
-	Political issue	<< Ecological–Materialist >>	Subsidies & environment	13. Nästa regering ska avveckla största delen av de miljöskadliga företagsstöden
-	Political issue	<< Left–Right >>	Pensions	14. De lägsta pensionerna ska höjas
-	Political issue	<< Left–Right >>	Collective agreement	15. Allmänbindande kollektivavtal ska slopas på sikt
-	Political issue	<< Conservative–Liberal >>	Criminal punishment	16. För de grövsta brotten ska livstids fängelse bokstavligen innebära livstid

-	Political issue	<< Conservative–Liberal >>	Sexual assault	17. Finland ska ha en samtyckeslag
-	Political issue	<< Immigration >>	Refugees	18. Asylansökningar ska behandlas utanför Europas gränser
Quantification	Political issue	<< Immigration >>	Refugees	19. Finland ska öka flyktingkvoten till minst 2 000 personer per år
-	Political issue	<< Immigration >>	Citizenship	20. Medborgarskap ska förutsätta ett test i medborgarkunskap och ett språkprov
Quantifications	Political issue	<< National–International >>	Development aid	21. Utvecklingsbiståndet ska höjas till 0,7 procent av bnp senast 2023
-	Political issue	<< Conservative–Liberal >>	Equality (gender)	22. Minst fyra månader föräldraledighet ska vikas för pappan eller den andra föräldern
-	Political issue	<< Conservative–Liberal >>	Equality (gender)	23. Finland ska införa obligatoriskt uppbåd för män och kvinnor
-	Political issue	<< Monolingual–Bilingual >>	Swedish language	24. Obligatorisk svenska ska återinföras i studentskrivningarna
-	Political issue	<< Left–Right >>	Elderly care	25. Vinster i skattefinansierad vård ska förbjudas

Table A3. Yle (2019) VAA statements

Problem	Type	Ideological dimension	Topic	Statement
-	Political issue	<< National–International >>	Euro	Euron ulkopuolella Suomi pärjäisi paremmin.
Double-barrelled	Political issue	Other	Higher education	Korkeakoulujen määrää pitää vähentää ja vapautuneet voimavarat käyttää huippuopetukseen ja -tutkimukseen.
-	Political issue	Other	Holidays	Koulujen kesälomia tulee siirtää kahdella viikolla niin, että ne alkavat kesäkuun puolivälissä ja päättyvät elokuun lopulla.
-	Political issue	<< Left–Right >>	Economy	Kun valtion menoja ja tuloja tasapainotetaan, se on tehtävä mieluummin menoja karsimalla kuin veroja kiristämällä.
Quantification	Political issue	<< Immigration >>	Immigration	Maahanmuuttajien määrän kasvu on lisännyt turvattomuutta Suomessa.
Quantification	Political issue	<< Ecological–Materialist >>	Forests	Metsiä hakataan Suomessa liikaa.
-	Political issue	<< Left–Right >>	Nato membership	Nato-jäsenyys vahvistaisi Suomen turvallisuuspoliittista asemaa.
Ambiguous	Opinion to non-issue	Other	Hierarchies	On oikein, että yhteiskunnassa jotkut ryhmät ovat paremmassa asemassa kuin toiset.

-	Political issue	Other	Education	Oppivelvollisuus pitää ulottaa myös ammatilliseen koulutukseen ja lukioon.
-	Political issue	<< Conservative–Liberal >>	Euthanasia	Parantumattomasti sairaalla on oltava oikeus eutanasiaan.
-	Political issue	<< Conservative–Liberal >>	Equality (gender)	Perhevapaita pitää uudistaa niin, että vapaat jakautuvat tasan vanhempien kesken.
Ambiguous	Personal value question	<< Conservative–Liberal >>	Traditional values	Perinteiset arvot ovat hyvän elämän perusta.
Ambiguous	Opinion to non-issue	Other	Politician's responsibility	Poliitikon velvollisuus on ennen kaikkea ajaa omien äänestäjiensä etuja.
-	Political issue	<< Immigration >>	Immigration & social services	Sosiaali- ja terveyspalveluiden rahoittaminen vaatii työperäisen maahanmuuton merkittävää lisäämistä.
-	Political issue	<< Left–Right >>	Health care	Sosiaali- ja terveyspalvelut on tuotettava ensisijaisesti julkisina palveluina.
Ambiguous	Political issue	<< Left–Right >>	Social security (basic income)	Sosiaaliturvaa tulee kehittää niin, että osa nykyisistä tuista korvataan kaikille työikäisille maksettavalla, vastikkeettomalla perustulolla.
-	Political issue	<< Conservative–Liberal >>	Sexual minorities	Sukupuolen korjaamisen tulee olla mahdollista myös alle 18-vuotiaille.
-	Political issue	<< Ecological–Materialist >>	Traffic emissions	Suomen ei pidä kiirehtiä kieltämään uusien bensa- ja dieselautojen myyntiä.
Ambiguous	Political issue	<< Conservative–Liberal >>	Individual responsibility	Suomen lakien pitäisi nykyistä vapaammin antaa ihmisten tehdä omat ratkaisunsa ja kantaa niiden seuraukset.
-	Political issue	<< Ecological–Materialist >>	Climate change	Suomen pitää olla edelläkävijä ilmastonmuutoksen vastaisessa taistelussa, vaikka se aiheuttaisi suomalaisille kustannuksia.
Ambiguous	Political issue	<< Conservative–Liberal >>	Law and Order	Suomessa tarvitaan nyt koviakin keinoja järjestyksen ja tavallisten ihmisten puolustamiseksi.
-	Political issue	<< Ecological–Materialist >>	Meat consumption	Valtion pitää ohjata suomalaiset syömään vähemmän lihaa esimerkiksi verotuksen avulla.
-	Political issue	<< Left–Right >>	Outsource elderly care	Vanhustenhoidon ulkoistamista yksityisille toimijoille tulee lisätä.
-	Political issue	Other	Hate speech	Vihapuhe tulee määritellä ja asettaa rangaistavaksi rikoslaissa.
-	Political issue	<< Conservative–Liberal >>	Alcohol	Viinit ja vahvat oluet pitää saada ruokakauppoihin.

Table A4. Helsingin Sanomat (2019) VAA statements

Problem	Type	Ideological dimension	Topic	Statement
-	Political issue	Other	Health care	Terveyskeskuksen lääkäri- ja hoitajakäyntien tulisi olla asiakkaalle maksuttomia.
Ambiguous	Political issue	<< Left–Right >>	Elderly care	Nykyään vanhuksen tulot vaikuttavat laitoshoidon hoitomaksuihin. Jatkossa myös vanhuksen omaisuutta tulisi käyttää hoivamaksujen kattamiseen.
-	Political issue	<< Left–Right >>	Social security	Suomen sosiaaliturvassa olisi syytä siirtyä kansalaispalkan eli vastikkeettoman sosiaaliturvan suuntaan.
-	Political issue	<< Left–Right >>	Social security	Sosiaaliturvaa tulisi uudistaa siten, että vastikkeeksi tuesta täytyy tehdä nykyistä enemmän jotakin yhteiskunnallisesti hyödyllistä, kuten opintoja tai vapaaehtoistöitä.
-	Political issue	<< Center–Periphery >>	Regional equality	Korkeakoulutusta on oltava tarjolla joka puolella Suomea.
-	Opinion to non-issue	Other	Childcare	Päiväkotien varhaiskasvatuksen tavoite on ensisijaisesti mahdollistaa vanhempien työssäkäynti.
-	Political issue	<< Conservative–Liberal >>	Equality (gender)	Kun perhevapaita uudistetaan, tärkeä tavoite on kasvattaa vain isälle suunnattua kiintiötä.
-	Political issue	<< Conservative–Liberal >>	Home care subsidy	Kotihoidon tukea ei saa lyhentää nykyisestä lapsen kolmesta ikävuodesta.
-	Political issue	<< Ecological–Materialist >>	Carbon budget	Suomessa tulee asettaa kansalaisille henkilökohtainen hiilibudjetti, jossa määritellään hiilijalanjäljelle enimmäisraja.
-	Political issue	<< Ecological–Materialist >>	Nuclear power	Eduskunnan pitäisi antaa lupia uusille ydinvoimaloille.
-	Political issue	<< Ecological–Materialist >>	Forests	Suomen pitää vähentää hakkuita ilmastonmuutoksen hillitsemiseksi.
-	Political issue	<< Ecological–Materialist >>	Meat consumption	Suomalaisia pitäisi ohjata vähäisempään lihansyöntiin, esimerkiksi verotuksen keinoin.
Quantification	Political issue	<< National–International >>	EU	EU:sta on Suomelle enemmän hyötyä kuin haittaa.
-	Political issue	<< Immigration >>	Multiculturalism	Suomen muuttuminen aiempaa monikulttuurisemmaksi ja monimuotoisemmaksi on hyvä asia.
-	Political issue	<< Immigration >>	Foreign students	Ulkomailta tuleville opiskelijoille pitäisi myöntää oleskelulupa koko tutkinnon suorittamisen ajaksi.
-	Political issue	<< Left–Right >>	Nato membership	Suomen tulisi tällä vaalikaudella ryhtyä valmistelemaan hakemista Natoon.
-	Political issue	Other	Job market reforms	Eduskunnan pitäisi päättää työmarkkinoita koskevista uudistuksista myös vastoin työmarkkinajärjestöjen eli ammattiliittojen ja työnantajajärjestöjen tahtoa.

-	Political issue	Other	Cars	Autoilu on Suomessa jo liian kallista.
-	Hypothetical question	Other	Robot cars	On oikein rajoittaa ihmisten autoilua, jos robottiautot ovat liikenteessä turvallisempia.
-	Political issue	Other	Medication sales rights	Päänsärkylääkkeitä ja vastaavia reseptittömiä itsehoitolääkkeitä pitäisi voida ostaa ruokakaupasta.
-	Political issue	<< Conservative–Liberal >>	Sexual minorities	Homo- ja lesbopareilla pitää olla samat avioliitto- ja adoptio-oikeudet kuin heteropareilla.
-	Political issue	<< Immigration >>	Immigration	Jos valtio tarjoaa turvapaikanhakijoiden vastaanottokeskuksen perustamista kotikuntaani, tarjous pitää hyväksyä.
Double-barrelled	Personal value question	<< Conservative–Liberal >>	School discipline	Kouluissa kohdellaan koululaisia liian lepsusti. Tiukempi kuri tekisi kouluista parempia.
-	Personal value question	<< Conservative–Liberal >>	Traditional values	Perinteiset arvot — kuten koti, uskonto ja isänmaa — muodostavat hyvän arvopohjan politiikalle.
-	Political issue	<< Left–Right >>	Outsourcing	Julkisia palveluita tulisi ulkoistaa entistä enemmän yksityisten yritysten tuotettavaksi.
-	Hypothetical question	<< Left–Right >>	Economy	Jos tulee eteen tilanne, jossa on välttämätöntä joko leikata julkisia palveluita ja sosiaalietuuksia tai korottaa veroja, veronkorotukset ovat parempi vaihtoehto.
-	Personal value question	<< Left–Right >>	Equality (income)	Suuret tuloerot ovat hyväksyttäviä, jotta erot ihmisten lahjakkuudessa ja ahkeruudessa voidaan palkita.
-	Political issue	<< Left–Right >>	Social services	Nykyisen kaltaiset palvelut ja sosiaalietuudet ovat pitkällä aikavälillä liian raskaita julkiselle taloudelle.
-	Political issue	<< Ecological–Materialist >>	Environment and economy	Talouskasvu ja työpaikkojen luominen tulisi asettaa ympäristöasioiden edelle silloin, kun nämä kaksi ovat keskenään ristiriidassa.
Double-barrelled	Political issue	<< Ecological–Materialist >>	Environment and economy	Kaikessa päätöksenteossa pitäisi arvioida vaikutukset ympäristöön ja tarvittaessa luopua ympäristölle haitallisista hankkeista.

Table A5. MTV Uutiset (2019) VAA statements

Problem	Type	Ideological dimension	Topic	Statement
-	Political issue	<< Left–Right >>	Taxation	2. Kokonaisveroastetta pitää edelleen laskea
-	Political issue	<< Left–Right >>	Social security (basic income)	3. Kun sosiaaliturvaa uudistetaan, käyttöön pitää ottaa uusi vastikkeeton perustulo
-	Political issue	Other	Regional taxation	4. Maakunnille pitää antaa verotusoikeus sote- ja maakuntauudistuksen valmistuttua
-	Political issue	<< Left–Right >>	Health care privatization	5. Sote-uudistusta on valmisteltu liiaksi yksityisten terveysyritysten ehdoilla.
-	Political issue	Other	Corporate subsidies	6. Teollisuuden saamia yritystukia voidaan karsia ilman, että yritysten kilpailukyky vaarantuu
-	Political issue	<< Ecological–Materialist >>	Forests	7. Suomen pitää vähentää metsien hakkuita ilmastonmuutoksen hillitsemiseksi.
-	Political issue	<< Ecological–Materialist >>	Flight tax	8. Suomessa pitää ottaa käyttöön lentomakustajilta perittävä lentovero.
-	Political issue	<< Ecological–Materialist >>	Nuclear energy	9. Ydinvoimaa tarvitaan energiantuotannossa ilmaston lämpenemisen hillitsemiseksi.
-	Political issue	<< Ecological–Materialist >>	Climate change	10. Ilmastonmuutoksen uhkia liioitellaan.
-	Political issue	<< Ecological–Materialist >>	Electric car subsidies	11. Sähköauton ostajan pitäisi saada valtion tukea hankintaansa.
-	Political issue	<< Ecological–Materialist >>	Gas motor ban	12. Seuraavan eduskunnan pitää tehdä päätös polttomoottoriautojen kieltämiseksi tulevien vuosikymmenten aikana.
-	Political issue	<< Left–Right >>	Public sector worker rights	13. Julkisen sektorin työntekijöiden lomaetuuksia voidaan leikata.
-	Political issue	<< Left–Right >>	Local work agreements	14. Työehdot ja palkat pitää sopia paikallisesti ilman liittoja ja keskusjärjestöjä.
-	Political issue	<< Left–Right >>	Unemployment benefit	15. Työttömyysturvaa voidaan leikata, jos työtön ei osoita aktiivisuutta työn tai koulutuksen hakemisessa.
-	Political issue	<< Left–Right >>	Worker rights	16. Irtisanomista pienissä alle 20 hengen yrityksissä pitää helpottaa.
-	Political issue	<< Immigration >>	Refugees	17. Turvapaikan hakemisen pitää tapahtua EU:n rajojen ulkpuolella sijaitsevissa keskuksissa.
-	Political issue	<< Immigration >>	Welfare chauvinism	18. Perustuslakia pitää muuttaa niin, että oleskeluluvan saaneilla turvapaikanhakijoille voidaan maksaa alempaaa sosiaalturvaa kuin Suomen kansalaisille.
Quantification	Political issue	<< Immigration >>	Refugees	19. Kiintiöpakolaisten määrää pitää lisätä nykyisestä 750:stä.

-	Political issue	<< Immigration >>	Refugees	20. Suomen pitää voida itse päättää, mistä maasta se ottaa pakolaisia humanitaarisin perustein.
-	Political issue	<< Immigration >>	Deportations	21. Rikokseen syyllistyneiden turvapaikanhakijoiden karkottaminen pitää tehdä nykyistä helpommaksi.
-	Political issue	<< Left–Right >>	Nato membership	22. Suomen pitää liittyä Naton jäseneksi.
-	Political issue	Other	Air defence	23. Suomen pitää valtion taloustilanteesta riippumatta korvata Hornet-hävittäjät täysimääräisesti eli hankkia 64 konetta.
-	Political issue	Other	Russia foreing policy	24. Venäjä on vastuussa Euroopan turvallisuuspoliittisen tasapainon järkkymisestä.
Quantification	Political issue	<< Conservative–Liberal >>	Police	25. Poliisien määrää pitää lisätä.
-	Political issue	Other	Extra child benefit	26. Syntyvyyttä on mahdollista lisätä maksamalla syntyvästä lapsesta kertaluontoinen rahallinen korvaus nykyisen lapsilisän lisäksi.
-	Political issue	<< Conservative–Liberal >>	Equality (gender)	27. Lakeja pitää muutta niin, että perhevapaat jakautuvat tasan molempien vanhempien kesken.
Double-barrelled	Political issue	<< Conservative–Liberal >>	Equality (gender)	28. Naisten työllisyysastetta pitää nostaa kotihoidontuen kestoa lyhentämällä.
-	Political issue	Other	Compulsory education age	29. Oppivelvollisuusikää pitää pidentää kattamaan myös toisen asteen koulutus
-	Political issue	Other	Preschool	30. Viisivuotiaat pitää velvoittaa esiopetukseen.
-	Political issue	Other	Elderly care	31. Ympärivuorokautisessa vanhustenhoidossa yhtä vanhusta kohden tarvittava hoitajamäärä pitää kirjata lakiin.
Ambiguous	Opinion to non-issue	<< Conservative–Liberal >>	Gender neutrality	32. Koulussa ja päiväkodeissa lapsia pitää puhutella sukupuolineutraalisti eikä tyttöinä ja poikina.
-	Political issue	<< Conservative–Liberal >>	Traditional values	33. Suvivirsi kuuluu koulun päättäjäisiin Suomessa.
-	Political issue	<< Conservative–Liberal >>	Alcohol	34. Viinit pitää saada ruokakauppaan seuraavien neljän vuoden kuluessa.
-	Political issue	<< Conservative–Liberal >>	Cannabis	35. Kannabiksen käyttö pitää sallia Suomessa.
-	Political issue	<< Ecological–Materialist >>	Meat consumption	36. Punaisen lihan syöntiä pitää rajoittaa ylimääräisellä verolla.

B. **Network visualizations of VAA results**

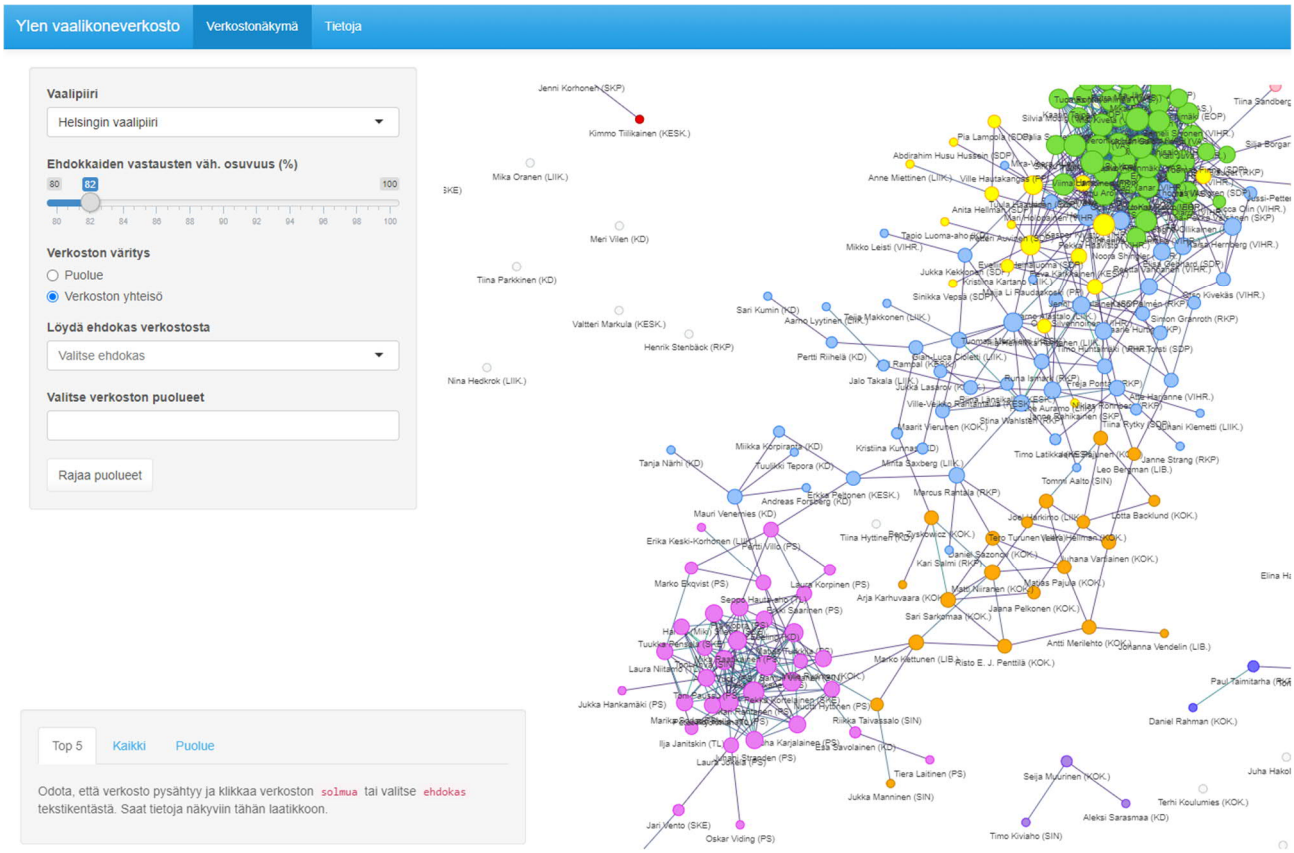

Figure B1. Community detection algorithm of the whole network, ties formed when matching score at least 82% (Ylen vaalikoneverkosto 2019)

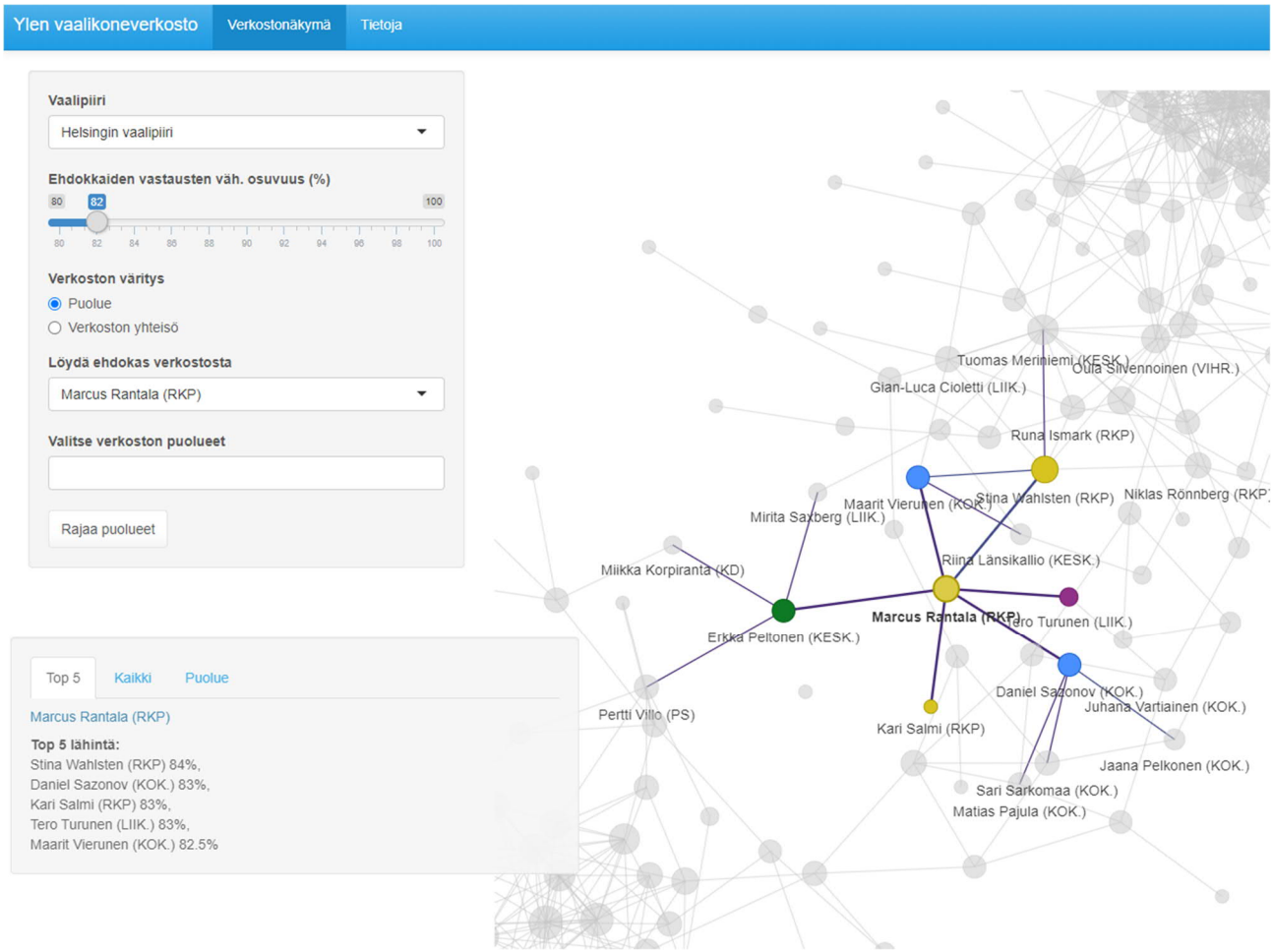

Figure B2. Candidate view of the network (candidate and directly linked nodes highlighted), showing top 5 closest matches to the candidate (Ylen vaalikoneverkosto 2019)