



cogitatio

POLITICS AND GOVERNANCE

Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives

Volume 14

2026

Open Access Journal

ISSN: 2183-2463

Edited by Diego Garzia, Stefan Marschall,
Mathias Wessel Tromborg, and Andreas Albertsen



Politics and Governance, 2026, Volume 14

Voting Advice Applications: Methodological Innovations, Behavioural Effects,
and Research Perspectives

Published by Cogitatio Press

Rua Fialho de Almeida 14, 2º Esq.,

1070-129 Lisbon

Portugal

Design by Typografia®

<http://www.typografia.pt/en/>

Cover image: © Designed by Freepik (www.freepik.com)

Academic Editors

Diego Garzia (University of Lausanne / Bologna University)

Stefan Marschall (Heinrich Heine University Düsseldorf)

Mathias Wessel Tromborg (Aarhus University)

Andreas Albertsen (Aarhus University)

Available online at: www.cogitatiopress.com/politicsandgovernance

This issue is licensed under a Creative Commons Attribution 4.0 International License (CC BY). Articles may be reproduced provided that credit is given to the original and *Politics and Governance* is acknowledged as the original venue of publication.

Table of Contents

Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives

Diego Garzia, Stefan Marschall, Mathias Wessel Tromborg, and Andreas Albertsen

Low Education, Low Impact? The Effects of Voting Advice Applications on an Underexposed Segment of Users

L. Constantin Wurthmann, Daniel C. Hagemann, and Stefan Marschall

Televised Influence: Examining Opinion Formation Through Live Completion of a Voting Advice Application

Nikandros Ioannidis and Vasiliki Triga

Voting Advice Application Use Increases Party Position Knowledge: An Experimental Study Among Belgian Youngsters

Joke Matthieu, Laura Jacobs, Matthias Van Campenhout, and Stefaan Walgrave

Are Single-Party Voting Advice Applications Useful? Comparing Voter Preferences in the BSW-O-Mat With a Probability Sample

Jan Philipp Thomeczek

Voting Advice Applications and Their Impact on Ideological and Affective Polarization

Maxime Walder, Jan Fivaz, and Daniel Schwarz

How Many People Use Voting Advice Applications? Survey Evidence From 15 Democracies Worldwide

Diego Garzia and Frederico Ferreira da Silva

Exploring Their Options? Tracking How Voters Actually Use Voting Advice Applications

Andreas Videbæk Jensen, Veikko Isotalo, and Mathias Wessel Tromborg

Enhancing Voting Advice Applications: Politicians' Perspectives on Additional Contextual Information and AI Integration

Elke van Veggel, Naomi Kamoen, and Christine Liebrecht

Estimating the Recommendation Certainty in Candidate-Based Voting Advice Applications

Fynn Bachmann, Daan van der Weijden, Cristina Sarasua, and Abraham Bernstein

In Pursuit of Informed Voters: Three Experimental Studies on Enhanced Voting Advice Applications

Naomi Kamoen, Christine Liebrecht, and Rieke van Lieshout

Table of Contents

Party Competition and Voter Attitudes in German Border Regions: Evidence From Local VAAs

Daniela Braun, Elisa Deiss-Helbig, Theresa Gessler, Jochen Müller, Julia Wagner,
and Georg Wenzelburger

Measuring Issue Congruence in Western Europe: How Voting Advice Applications Compare to Expert- and Manifesto-Based Estimates

Mattia Gatti

Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives

Diego Garzia ¹ , Stefan Marschall ² , Mathias Wessel Tromborg ³ ,
and Andreas Albertsen ³ 

¹ Department of Political and Social Sciences, University of Bologna, Italy

² Department of Social Sciences, Heinrich Heine University Düsseldorf, Germany

³ Department of Political Science, Aarhus University, Denmark

Correspondence: Diego Garzia (diego.garzia@unibo.it)

Submitted: 4 March 2026 **Published:** 26 March 2026

Issue: This editorial is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Voting advice applications (VAAs) are widely used digital tools that match voters’ policy preferences with party or candidate positions. Since their emergence in the 1980s and rapid expansion alongside internet dissemination, VAAs have spread to more than 80 countries and generated growing academic interest. This thematic issue advances the field by analysing behavioural effects, user diversity, design innovations, including AI-driven conversational agents, and the validity of VAA-generated data. Together, the contributions demonstrate both the maturation of VAA research and its continued relevance for understanding democratic representation.

Keywords

artificial intelligence; democratic representation; electoral behaviour; political knowledge; voting advice applications

1. Voting Advice Applications: Tools for Voters and Researchers

Voting advice applications (VAAs) have become increasingly widespread in electoral campaigns in democracies worldwide. These popular online tools are designed to help voters compare their policy preferences on major issues with the positions of political parties/candidates. Respondents fill out an online questionnaire with their opinion on a set of policy statements (e.g., “The state should provide stronger financial support to the unemployed”). The users’ answers are then compared with the position of the parties/candidates on the same

issues and offer a party or candidate “advice” to the user, for example, in the form of a rank-ordered list (for a detailed explanation of VAA operating principles and features, see Garzia & Marschall, 2019).

VAAAs can be traced back to the 1980s (Cedroni & Garzia, 2010), but their expansion geographically and in terms of usage numbers went hand in hand with the spread of internet connections. An example is the Dutch Stemwijzer, a pioneer of modern VAAAs that went from only 6,500 users in its first online edition of 1998 to over 9 million in 2023. The German Wahl-O-Mat has been available since 2002 and has been used over 150 million times since then. During the 2025 federal election alone, it was used around 26 million times.

According to the global census conducted in 2016 by the ECPR Research Network on Voting Advice Applications, VAAAs have been fielded in as many as 43 countries across five continents, and some countries have fielded multiple VAAAs simultaneously (Garzia & Marschall, 2016, 2019). Since then, VAAAs have been consolidated and introduced in new countries and contexts. According to a more recent census of September 2025, at least one VAA has been initiated in more than 80 countries across five continents (Marschall et al., 2026).

Due to their growing popularity, today VAAAs constitute a field of social science with large collaborative research projects and publications in leading academic journals. The last two decades have witnessed a growing interest in the consistency and reliability of the voting advice provided by these applications. Additionally, political scientists have shown particular interest in the effects of these tools on users’ electoral behaviour. Finally, VAAAs have become a valuable tool for collecting data to address key questions in democratic representation and party research across national and supranational electoral arenas.

The growing academic interest in VAAAs is illustrated in Figure 1, which plots the number of scientific publications mentioning VAAAs since 2003, calculated based on Google Scholar data.

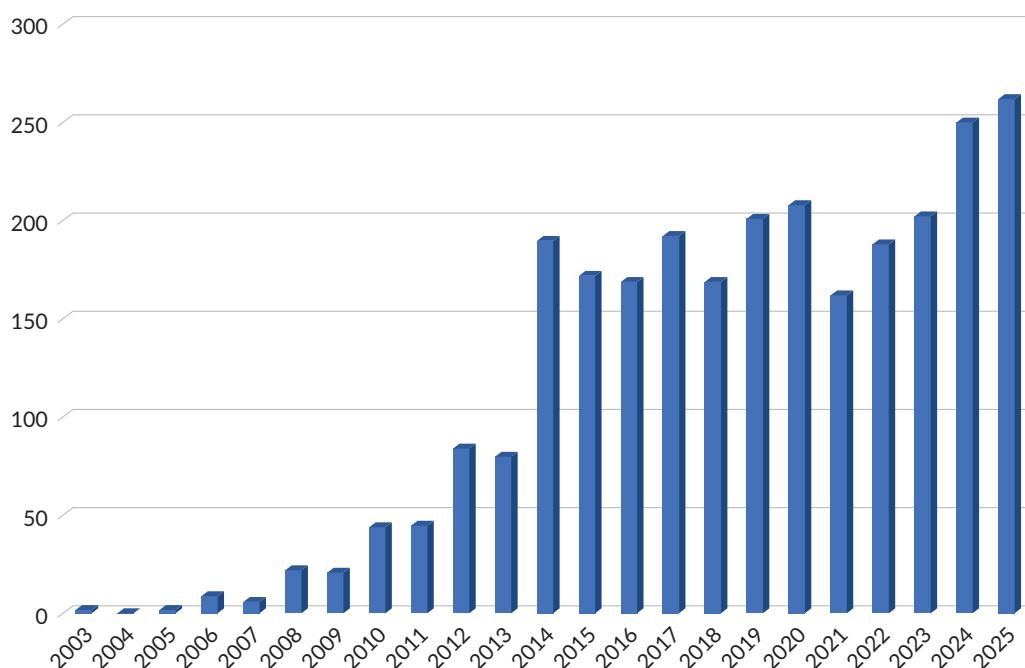


Figure 1. The number of academic publications referring to VAAAs, by year.

2. State of the Literature

Given the wide popularity of VAAs, the question of whether they affect political behaviour has represented a recurrent theme in the existing academic literature. By reducing the costs of information gathering, they have been hypothesized to stimulate political interest, communication, and knowledge among their users, thereby potentially increasing their likelihood of voting and/or switching party/candidate choice in line with the advice provided by the VAA platform. Munzert and Ramirez-Ruiz (2021) have delivered one of the most encompassing comparative assessments of VAA effects. Their meta-analysis of 22 studies published between 2006 and 2019 focuses on the influence of VAAs on turnout, party choice, and issue knowledge. The authors find evidence that VAAs affect both turnout and vote choice; however, the magnitude and statistical significance of the effects vary across empirical designs and contexts. Consequently, they call for higher-powered experimental studies to better assess these behavioural outcomes. More recent works provide further evidence of the link between VAA usage and political behaviour, especially regarding patterns of party/candidate choice (Benesch et al., 2023; Germann et al., 2023; Tromborg & Albertsen, 2023). Other findings paint a more mixed picture, finding, for example, no effect on turnout, but an effect on party choice (Frese et al., 2025), or finding no effect on turnout and vote choice, but an increase in knowledge about party positions (Munzert et al., 2020). Understanding how and when VAA advice affects voting behaviour remains a key part of the VAA research agenda.

A related theme of VAA research pertains to who uses them (Albertsen, 2022; Cedroni & Garzia, 2010; Garzia & Marschall, 2012) and how they use them (van de Pol et al., 2014). Much early research suggested that early users are more often politically interested and well educated. More recently, the focus has been on so far under-researched groups of users.

Another key question is whether the advice provided by VAAs is reliable and unbiased. The literature on VAA development and methodology can be summarized under four main aspects: the way the set of statements is selected and the statements are formulated (Lefevere & Walgrave, 2014); the method to position the parties with view to the statements (Ferreira da Silva et al., 2023; König & Nyhuis, 2020); the algorithm applied to calculate the proximity between the users' and parties' positions (Louwerse & Rosema, 2014); and the form in which the results are displayed (Bruinsma, 2020).

More recently, VAA practitioners and researchers have engaged with artificial intelligence and devoted attention to the so-called Conversational Agent VAAs (Gemenis, 2024; van Veggel et al., 2025; Velez et al., 2025). VAAs have also been used as a source of data on what voters and parties/politicians believe. The articles in this thematic issue make important contributions to the various aspects of the literature noted above.

3. Plan of the Thematic Issue

Several articles in this thematic issue develop the theme of how VAAs affect users. Based on data from the 2021 German Federal Election, Wurthmann et al. (2025) examine a specific group of users and demonstrate that formal education does not directly moderate VAA-driven mobilization or vote conversion. The analysis reveals that learning effects are central: users with lower or medium levels of education gain knowledge of party positions, which increases mobilization, reconsideration, or reinforcement of vote choice. Ioannidis and

Triga (2026) analyse how the televised use of a VAA by presidential candidates influenced voter opinion formation during the 2023 Cyprus elections. Using a natural experiment and regression discontinuity models, they find that candidate cues increased voter-candidate alignment for one candidate but produced mixed or even divergent effects for others. The results show that candidate cues can both foster convergence and expose divisions, depending on issue type, candidate profiles, and coalition heterogeneity. Matthieu et al. (2026) investigate in a large-scale experiment in Flanders, Belgium, whether VAAs improve young voters' knowledge of party positions. They find that exposure to a standard VAA did not increase party position knowledge, whereas a youth-oriented VAA significantly improved young respondents' ability to correctly identify party positions. The results highlight that VAAs can foster political learning when tailored to specific audiences. Thomeczek (2026) assesses whether data from a single-party VAA, the BSW-O-Mat, can be used to validly analyse voter preferences despite strong self-selection bias. By comparing VAA-based regression results with those from the German Longitudinal Election Study probability sample, the article finds substantial convergence in effect directions, though VAA data tend to overestimate effect sizes. The study concludes that single-party VAA data can still yield meaningful insights into voter preferences when biases are carefully accounted for. Walder et al. (2026) examine whether VAAs can reduce ideological and affective polarization among voters. Using data from the Swiss VAA Smartvote and a two-wave panel field experiment around the 2023 Swiss national election, they find partial evidence that VAA use is associated with decreased polarization. These results suggest that VAAs may play a depolarizing role in increasingly polarized democracies.

This thematic issue also offers new evidence regarding how VAAs are used and who their users are: In their short note, Garzia and Ferreira da Silva (2026) provide updated survey data on the use of VAAs across 15 democracies worldwide. Drawing on data from 48 national elections between 2003 and 2024, they show that VAA usage has steadily increased in most countries and that the characteristics of VAA users have evolved over time, indicating a changing user community. Jensen et al. (2026) examine how voters actually use VAAs by tracking real usage behaviour in a representative Danish sample. They find that while most users focus on their preferred party, many also explore alternative party options and spend considerable time comparing them. This behaviour is especially common among undecided and politically interested users.

Regarding VAA design, van Veggel et al. (2026) examine how VAAs can be enhanced with additional contextual information, such as clarifications, status quo explanations, debate summaries, and party positions, to improve users' understanding of political statements. Based on interviews with local Dutch politicians, they find broad support for adding these types of information. While politicians see potential in using artificial intelligence to deliver information, they stress that any AI integration must ensure accuracy, political neutrality, and transparency. Bachmann et al. (2026) introduce a method to estimate the certainty of recommendations in candidate-based VAAs when users do not complete all questions. They present an algorithm that accurately predicts candidate recommendation stability and outperforms heuristic approaches in simulations. Kamoen et al. (2026) examine whether VAAs with integrated chatbots (CAVAAs) help voters understand political statements better. Across three experimental studies, CAVAAs reduce user confusion and increase perceived knowledge and usability compared to standard VAAs, with effects consistent across different levels of political sophistication.

Finally, this thematic issue provides clear examples of how VAA data can be used to explore important questions in political science. Braun et al. (2026) address party competition and voter attitudes toward

Europe in EU border regions, where local politics often closely intersect with European integration. Using data from the VOTO, a VAA for the 2024 German local elections, they analyse party positions and citizens' preferences in four German border regions. The findings show that border-specific contexts and party cues shape distinctive voter attitudes on cross-border cooperation and European integration, supporting both qualitative and quantitative research on border-region political behaviour. Employing data from VAAs requires, of course, that such data is reliable. This is the topic of another contribution in this issue. Gatti (2026) examines party-voter congruence beyond ideological dimensions, highlighting persistent problems with missing and inconsistent measures of perceived party positions. The study cross-validates traditional survey-based indicators with VAA-derived data and finds broad convergence between the approaches. Overall, the results support the validity of VAA-generated data.

While early debates in VAA research primarily addressed fundamental and foundational questions, the contributions to this thematic issue indicate that the field has matured to a stage at which in-depth analyses are both possible and necessary. The articles clearly demonstrate the level of theoretical and methodological specification and reflection at which current debates in VAA research are conducted. At the same time, changing contextual conditions, most notably the growing role of artificial intelligence, have generated new research questions, underscoring that substantial work remains to be done to fully understand the role of VAAs in democracies.

Conflict of Interests

The authors declare no conflict of interests.

References

- Albertsen, A. (2022). How do the characteristics of voting advice application users change over time? Evidence from the German election studies. *German Politics*, 31(3), 399–419.
- Bachmann, F., van der Weijden, D., Sarasua, C., & Bernstein, A. (2026). Estimating the recommendation certainty in candidate-based voting advice applications. *Politics and Governance*, 14, Article 11256. <https://doi.org/10.17645/pag.11256>
- Benesch, C., Heim, R., Schelker, M., & Schmid, L. (2023). Do voting advice applications change political behavior? *The Journal of Politics*, 85(2), 684–700.
- Braun, D., Deiss-Helbig, E., Gessler, T., Müller, J., Wagner, J., & Wenzelburger, G. (2026). Party competition and voter attitudes in German border regions: Evidence from local VAAs. *Politics and Governance*, 14, Article 11228. <https://doi.org/10.17645/pag.11228>
- Bruinsma, B. (2020). Evaluating visualisations in voting advice applications. *Statistics, Politics and Policy*, 11(1), 1–21.
- Cedroni, L., & Garzia, D. (2010). *Voting advice applications in Europe: The state of the art*. Civis.
- Ferreira da Silva, F., Reiljan, A., Cicchi, L., Trechsel, A. H., & Garzia, D. (2023). Three sides of the same coin? Comparing party positions in VAAs, expert surveys and manifesto data. *Journal of European Public Policy*, 30(1), 150–173.
- Frese, J., Hix, S., & Lachat, R. (2025). Quality not quantity: How a VAA affected voting behavior in three large-scale field experiments. *British Journal of Political Science*, 55, Article e176.
- Garzia, D., & Ferreira da Silva, F. (2026). How many people use voting advice applications? Survey evidence from 15 democracies worldwide. *Politics and Governance*, 14, Article 11347. <https://doi.org/10.17645/pag.11347>

- Garzia, D., & Marschall, S. (2012). Voting advice applications under review: The state of research. *International Journal of Electronic Governance*, 5(3/4), 203–222.
- Garzia, D., & Marschall, S. (2016). Research on voting advice applications: State of the art and future directions. *Policy & Internet*, 8(4), 376–390.
- Garzia, D., & Marschall, S. (2019). Voting advice applications. In W. Thompson (Ed.), *Oxford Research Encyclopedia of Politics*. Oxford University Press.
- Garzia, D., Marschall, S., Tromborg, M. W., & Albertsen, A. (2026). Voting advice applications: Methodological innovations, behavioural effects, and research perspectives. *Politics and Governance*, 14, Article 12331. <https://doi.org/10.17645/pag.12331>
- Gatti, M. (2026). Measuring issue congruence in Western Europe: How voting advice applications compare to expert- and manifesto-based estimates. *Politics and Governance*, 14, Article 11259. <https://doi.org/10.17645/pag.11259>
- Gemenis, K. (2024). Artificial intelligence and voting advice applications. *Frontiers in Political Science*, 6, Article 1286893.
- Germann, M., Mendez, F., & Gemenis, K. (2023). Do voting advice applications affect party preferences? Evidence from field experiments in five European countries. *Political Communication*, 40(5), 596–614.
- Ioannidis, N., & Triga, V. (2026). Televised influence: Examining opinion formation through live completion of a voting advice application. *Politics and Governance*, 14, Article 11272. <https://doi.org/10.17645/pag.11272>
- Jensen, A. V., Isotalo, V., & Tromborg, M. W. (2026). Exploring their options? Tracking how voters actually use voting advice applications. *Politics and Governance*, 14, Article 11277. <https://doi.org/10.17645/pag.11277>
- Kamoen, N., Liebrecht, C., & van Lieshout, R. (2026). In pursuit of informed voters: Three experimental studies on enhanced voting advice applications. *Politics and Governance*, 14, Article 11234. <https://doi.org/10.17645/pag.11234>
- König, P. D., & Nyhuis, D. (2020). Assessing the applicability of vote advice applications for estimating party positions. *Party Politics*, 26(4), 448–458.
- 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.
- Louwerse, T., & Rosema, M. (2014). The design effects of voting advice applications: Comparing methods of calculating matches. *Acta Politica*, 49, 286–312.
- Marschall, S., Hagemann, D., & Bongartz, J. (2026). *Der Wahl-O-Mat*. Springer.
- Matthieu, J., Jacobs, L., Van Campenhout, M., & Walgrave, S. (2026). Voting advice application use increases party position knowledge: An experimental study among Belgian youngsters. *Politics and Governance*, 14, Article 11235. <https://doi.org/10.17645/pag.11235>
- Munzert, S., Barbera, P., Guess, A., & Yang, J. (2020). Do online voter guides empower citizens? Evidence from a field experiment with digital trace data. *Public Opinion Quarterly*, 84(3), 675–698.
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706.
- Thomeczek, J. (2026). Are single-party voting advice applications useful? Comparing voter preferences in the BSW-O-Mat with a probability sample. *Politics and Governance*, 14, Article 11241. <https://doi.org/10.17645/pag.11241>
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15(4), 582–599.
- 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.

- van Veggel, E., Kamoen, N., & Liebrecht, C. (2026). Enhancing voting advice applications: Politicians' perspectives on additional contextual information and AI integration. *Politics and Governance*, 14, Article 11146. <https://doi.org/10.17645/pag.11146>
- van Veggel, E., Liebrecht, C., & Kamoen, N. (2025). How life-like digital humans in voting advice applications can stimulate young voters to inform themselves about politics: The role of familiarity and expertise. *International Journal of Human-Computer Interaction*, 41(20), 12816–12830.
- Velez, Y. R., Green, D. P., & Sevi, S. (2025). Chatbot voting advice applications inform but seldom sway young unaligned voters. *Proceedings of the National Academy of Sciences*, 122(50), Article e2515516122.
- Walder, M., Fivaz, J., & Schwarz, D. (2026). Voting advice applications and their impact on ideological and affective polarization. *Politics and Governance*, 14, Article 11236. <https://doi.org/10.17645/pag.11236>
- Wurthmann, L., Hagemann, D., & Marschall, S. (2025). Low education, low impact? the effects of voting advice applications on an underexposed segment of users. *Politics and Governance*, 14, Article 11087. <https://doi.org/10.17645/pag.11087>

About the Authors

Diego Garzia is an associate professor of political science at the University of Bologna, Italy. His research interests focus on elections, public opinion, and voting behavior in a comparative perspective. He currently serves in the Steering Committees of the Consortium of National Election Studies and the ECPR Research Network on Voting Advice Applications.

Stefan Marschall is a full professor of political science at the Heinrich-Heine-University of Düsseldorf, Germany. He is a specialist on the political system of Germany, on comparative as well as transnational parliamentarism, and especially on political online communication and participation. He serves as a member of the Steering Committee of the ECPR Research Network on Voting Advice Applications.

Mathias Wessel Tromborg is an associate professor of political science at Aarhus University. His research lies at the intersection of voter behavior, politician behavior, and political institutions. He is currently the Principal Investigator (PI) of the research project The Role of Voting Advice Applications in Representative Democracy. He serves as a member of the Steering Committee of the ECPR Research Network on Voting Advice Applications.

Andreas Albertsen is an associate professor of political science at Aarhus University. His research in political theory has a particular interest in the ethics of democratic participation. He serves as a member of the Steering Committee of the ECPR Research Network on Voting Advice Applications.

Low Education, Low Impact? The Effects of Voting Advice Applications on an Underexposed Segment of Users

L. Constantin Wurthmann ¹ , Daniel C. Hagemann ² , and Stefan Marschall ² 

¹ Mannheim Centre for European Social Research (MZES), University of Mannheim, Germany

² Department of Social Sciences, Heinrich Heine University Düsseldorf (HHU), Germany

Correspondence: Daniel C. Hagemann (daniel.hagemann@hhu.de)

Submitted: 1 August 2025 **Accepted:** 27 October 2025 **Published:** 10 December 2025

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Voting Advice Applications (VAAs) play a crucial role in the democratic pre-election period, positively affecting political interest, voter participation, and political knowledge. A number of studies show that they are predominantly used by people with formal higher education levels and a high level of political interest. This article takes this as a starting point to examine the mechanisms that lead individuals to feel motivated to participate in an election, feel reinforced in their voting decision, or even be converted to vote for another party through a VAA. Using data collected before the 2021 German Federal Election and drawing on path models, we find that formal education does not directly impact VAA-induced electoral mobilization or vote conversion at first glance. More specifically, we find a negative effect of average or lower formal education levels on being reinforced in an existing voting decision by the Wahl-O-Mat. However, an analysis of the mechanisms reveals that learning effects attributed to the Wahl-O-Mat usage play a key role. We observe that individuals with medium or low education gain increased awareness of party positions and differences. This, in turn, significantly raises the likelihood of being mobilised, reconsidering party choice, or reinforcing an existing decision. The findings are highly relevant for VAA research as they examine the specific preconditions and mechanisms of the effects of such a tool. For individuals with lower education levels, the learning experience through VAAs seems to be particularly crucial. For the development of VAAs, the findings show how crucial the respective learning environment is for the user experience.

Keywords

education; knowledge; political sociology; public opinion; voting; voting behaviour

1. Introduction

Over the past two decades, a new type of online tool has proliferated in European democracies and beyond: Voting advice applications (VAAs). These applications assist users in casting their votes by aligning their policy preferences with the programmatic stances of political parties or candidates running for election (Garzia & Marschall, 2016). Users complete a questionnaire, expressing their views on a series of policy proposals selected for their relevance to the current campaign. The users' responses are then compared to the positions of the parties or candidates on these same issues. "Advice" is subsequently generated, typically presented as a rank-ordered list. By providing a ranking of viable options, VAAs clarify the political landscape for users, highlighting how their preferences align with the competing political entities. In many countries, VAAs have become obligatory elements in pre-election times on which large shares of the electorate draw.

The research on the effects of VAAs has a long and established tradition (Garzia & Marschall, 2019). A key focus of this body of work has been the "preaching to the converted" thesis, which suggests that VAAs primarily reach out to those who are already politically engaged and those who are formally highly educated (Albertsen, 2022; Marschall, 2014). Given that education is a key determinant of political participation and interest (Smets & van Ham, 2013; Verba et al., 1995), this highlights the aspect of a selective reach of VAAs and asks the question of how individuals with different educational backgrounds engage with these tools. Studies in VAA research have predominantly worked with representative user samples or experimental groups dominated by individuals who are already mobilised, with the tendency to focus on those who are not yet engaged in the political process (Munzert & Ramirez-Ruiz, 2021). This approach limits the understanding of VAAs' potentially differentiated impact on voter engagement and participation and ignores their potential effects on different user groups.

There are significant studies discussing the strength and extent of VAAs' influence on voting decisions across different demographic and social groups (Germann et al., 2023; Tromborg & Albertsen, 2023). While some correlations have been established, the nuances of these effects remain unclear. The impact of VAAs may vary widely among different groups, but the precise mechanisms and character of these variations are not well understood.

It is essential to gain a deeper insight into how these tools function within various groups: The specific mechanisms through which VAAs influence voter behaviour are often described as a "black box" (Enyedi, 2015, p. 1000; van der Linden & Dufresne, 2017). We argue that learning processes play an important role when it comes to the effectiveness of VAAs. Then again, the extent to which users learn from these applications is likely influenced by their political sophistication. Users with different levels of political awareness and understanding may profit differently from interacting with VAAs. Understanding these learning dynamics is crucial for assessing how VAAs can be designed to be more effective for a broader audience.

Opening the "black box" requires a specific methodical approach: The bulk of research in this field has primarily relied on simple correlation analyses, which do not capture the complex interplay of factors that contribute to VAAs' effectiveness (Munzert & Ramirez-Ruiz, 2021). To gain a deeper understanding of how VAAs work, there is a need for more sophisticated methods, such as path analysis. Such an approach could uncover the intricate relationships between various factors influencing voter behaviour and the specific

pathways through which VAAs exert their effects. Moreover, by explicitly incorporating educational background as a moderating factor, we aim to shed light on whether VAAs reinforce or mitigate structural inequalities in political engagement. In short, while there is a substantial body of research on the effects of VAAs, significant research gaps remain. Focusing on the seemingly “non-convertible” individuals, clarifying the strength of VAA effects across different groups, and unpacking the “black box” of mechanisms through more complex analyses might represent essential steps forward. Additionally, examining the role of learning processes and how they are influenced by users’ prior knowledge will be critical for understanding the impact of VAAs on voter engagement and decision-making.

In this article, we begin by outlining the role of education as a key structural factor shaping political participation, voter engagement, and ideological alignment. We build on existing research that highlights education as a central determinant and explore its impact on the likelihood of using VAAs. Given that higher formal education levels are associated with greater political interest and sophistication, while lower levels are linked to political disengagement, we examine whether VAAs help bridge this divide or reinforce existing disparities. In doing so, we take the “preaching to the converted” hypothesis as a starting point, asking whether VAAs primarily benefit already mobilised, highly educated individuals or whether they provide a learning effect that is particularly relevant for less educated users.

In order to do so, we will draw on data collected before the 2021 German Federal Election; in the data analysis, we will be utilising path models to identify the mechanisms through which learning effects instigated by the most popular German VAA—the Wahl-O-Mat—operate. These findings suggest that the learning experience provided by VAAs is particularly important for individuals with medium or low formal education levels. For the development of VAAs, these results are significant, as they highlight the critical role of the learning environment provided by these tools.

2. Theory

2.1. *The Educational Divide?*

While socio-structural characteristics are often viewed primarily from the perspective of classical cleavages (Elff & Roßteutscher, 2022), education has developed into a “master structural variable” (Hooghe & Marks, 2025, p. 7). In particular, the highly educated represent a socio-structurally relevant group substantively concerned with political behaviour and social inequality. Overall, higher education attainment is influenced by a complex interplay of socioeconomic background, institutional structures, cultural values, and individual abilities. Families with higher socioeconomic status can provide academic support and instill long-term educational aspirations, making parental education and financial resources critical factors (Bourdieu & Passeron, 1990; Plutzer, 2002). Additionally, institutional mechanisms such as school tracking systems, government funding, and scholarship opportunities play a significant role in determining who can access tertiary education (Pfeffer, 2008). Beyond these structural elements, societal values regarding meritocracy, gender norms, and ethnic disparities further shape educational trajectories (Breen et al., 2009).

Given these mechanisms, it seems logical that education is a central determinant in explaining respondents’ political interests and knowledge (Hansen & Pedersen, 2014; Smets & van Ham, 2013). The higher the level of formal education, the more likely individuals are to participate in different forms of political engagement

(Verba et al., 1995). This is particularly evident in voter turnout, where a consistent effect of the education level can be observed over time, further reinforced by family socialisation (Dassonneville & Hooghe, 2017; Plutzer, 2002). Especially among younger cohorts, differences in formal education levels play a substantial role in predicting whether individuals participate in elections (Schäfer et al., 2020). Overall, the gap in voter turnout between citizens with low and high levels of education has expanded over time, particularly in countries experiencing sharper declines in electoral participation. The influence of education on the likelihood of voting has doubled, leading to an increasingly less representative electorate, with highly educated individuals significantly more likely to cast their ballots (Dassonneville & Hooghe, 2017). Dimensions of social inequality reflected in education levels thus become increasingly politically entrenched over time (Schäfer et al., 2020). These shared experiences not only foster a sense of group identity among individuals with similar levels of education (Westheuser & Zollinger, 2025). Education contributes to structuring social identities by shaping how individuals feel connected to specific groups. This effect is particularly pronounced along the universalist–particularist divide, which increasingly shapes European party systems. Individuals tend to feel a stronger affinity with others of a similar educational background, a pattern observed among highly and less-educated individuals (Bornschier et al., 2024). Additionally, de Jong and Kamphorst (2025, p. 1533) provide evidence that “large proportions of higher—and lower-educated citizens report no close relationships with individuals from different educational backgrounds.”

At the same time, education shapes the values and attitudes individuals adopt. Those with higher levels of formal education are more likely to embrace universalist, post-materialist, and cosmopolitan values, distinguishing themselves from those with lower levels of education, who are more inclined towards traditional, conservative, and materialist values (Elff & Roßteutscher, 2022; Hooghe & Marks, 2025; Marks et al., 2023). As a result, this structural characteristic undoubtedly influences political party competition (Abou-Chadi & Hix, 2021), being reinforced by educational disparities that progressively widen over successive generations (Schäfer & Steiner, 2025). For example, Stubager (2013) finds that an education-based cleavage has emerged as a key structuring force in Danish politics, driven by the increasing salience of authoritarian–libertarian values and reinforced by party competition, contrasting highly educated, libertarian-leaning voters with less-educated, authoritarian-leaning voters. More broadly, Bornschier et al. (2024, p. 30) find that “education is at the heart of cleavage formation,” as it interacts with group loyalties, identities, and in-group favouritism. However, the educational level achieved is not static; it can transform over time due to increased educational attainment. Higher education levels have a causal effect on reducing prejudice and fostering more liberal attitudes (Scott, 2022). It is, therefore, only logical that educational background can explain voting behaviour for certain parties and a lower likelihood of adopting populist attitudes (Häusermann & Kriesi, 2015; Koch et al., 2021).

At the same time, higher levels of formal education reduce strict party loyalty and significantly increase the consistency of individuals’ ideological positions (Nie et al., 1979). The highly educated are, therefore, more easily mobilised. Still, they are also more selective in their political choices and less bound by long-term determinants compared to those with lower levels of education (Elff & Roßteutscher, 2022). This can be explained by the fact that higher education levels promote political sophistication and critical thinking, making individuals more likely to question party loyalty rather than remain attached to a single party (Dalton, 2012). Mayer and Schultze (2019) show that more educated individuals are more likely to have multiple party identifications within a given political camp.

Politically critical and ideologically consistent, individuals with higher levels of education represent a segment of the electorate that is challenging to address through one-dimensional campaign strategies or to mobilise comprehensively. While they tend to process political information more quickly, the intensified flow of information during election campaigns can enable those with lower levels of education to somewhat reduce this informational gap (Hansen & Pedersen, 2014; Wurthmann & Marschall, 2022).

2.2. Overcoming the Educational Divide Using VAAs?

VAAs have increasingly assumed a central role in conveying party positions (Heinsohn et al., 2016), motivating electoral participation (Germann et al., 2023; Tromborg & Albertsen, 2023), and even fostering non-election-related forms of participation (Manavopoulos et al., 2018). In this section, we will examine the significance of VAAs in greater detail and outline why we assume that they also play a relevant role in addressing the educational divide.

VAAs are digital tools designed to help voters make informed electoral decisions by aligning their policy preferences with the positions of political parties or candidates. By answering a series of issue-based questions, users receive a personalised recommendation indicating which party or candidate best reflects their views. While primarily developed to encourage political engagement (Garzia, 2010; Garzia & Marschall, 2019; Wall et al., 2014), VAAs have become indispensable instruments in modern election campaigns, shaping voter behaviour and strategic party competition (Wurthmann & Marschall, 2022).

Widely used in European democracies, VAAs have demonstrable effects on voter turnout, political engagement, and party competition (Garzia, 2010; Germann et al., 2023; Marschall & Schultze, 2012; Wall et al., 2014). Research suggests that they help mitigate information asymmetries, particularly among individuals with lower levels of political knowledge or engagement (Kamoen et al., 2017; Tromborg & Albertsen, 2023). Moreover, VAAs have been shown to mobilise individuals who are younger (Germann & Gemenis, 2019), traditionally less engaged in politics, and unlikely to be reached by conventional party outreach efforts (Israel et al., 2017; Marschall & Schultze, 2012).

However, in line with the “uses and gratification theory,” the actual effects could be moderated by the concrete motivation of users to resort to VAAs. Van de Pol et al. (2014), for example, demonstrate that highly educated users report lower learning effects because they often employ VAAs primarily as confirmation tools rather than as instruments for knowledge acquisition. Similar dynamics are observable in more recent contexts. Evidence from Austrian VAAs in 2024 shows that a considerable share of users, both occasional and frequent, report that they “already know how they will vote” (Dieing, 2025, p. 9)—regardless of their educational backgrounds. Such findings indicate that only a subset of users is genuinely open to impacts made by VAAs, e.g., to preference change or political learning.

The extent to which VAA-impact varies—also based on education—remains an open question, as individuals with higher levels of formal education may process political information differently from those with lower levels. VAAs are crucial in promoting issue-based voting by clarifying party positions and helping voters navigate an increasingly complex political landscape. They provide users with structured access to party manifestos, facilitate reflection on key policy debates, and encourage deeper engagement with politics (Heinsohn et al., 2016; Kamoen et al., 2017; van de Pol, 2016).

VAA's allow users to familiarise themselves with political party positions, reflect on their own views, and develop a more informed electoral choice. Empirical findings indicate that VAA's can enhance political knowledge, particularly among those with lower prior political engagement levels (Germann et al., 2023; Mendez, 2017; Munzert et al., 2020; Stadelmann-Steffen et al., 2023). Additionally, Tromborg and Albertsen (2023) find that politically less engaged individuals without strong party identification are more likely to adjust their voting preferences after using VAA's. This effect is further supported by evidence that individuals with lower levels of prior political information are more susceptible to preference shifts following VAA usage (Germann et al., 2023). Such findings suggest that how highly and less educated voters respond to VAA's may stem from variations in political interest, media consumption, and prior knowledge levels. VAA's have also been linked to increased political media use, fostering broader political engagement (Manavopoulos et al., 2018).

While VAA's seek to make political information more accessible to all voters, their measured impact based on education remains inconclusive. Some studies suggest that VAA's effectively increase political knowledge across educational levels (Gemenis & Rosema, 2014; Heinsohn et al., 2016; Kamoen et al., 2017). However, it remains unclear how subjectively perceived learning effects influence political behaviour, particularly regarding the educational divide. Since higher-educated individuals tend to be more sceptical of politics in general, it is reasonable to assume that their perception of VAA's—either as a mere “gadget” or as a valuable tool for political education (Marschall & Israel, 2014)—differs significantly. Depending on respondents' educational levels and drawing on early research on election campaign effects (see, e.g., Lazarsfeld et al., 1944), as well as the growing importance of VAA's as an election tool in recent years (Wurthmann & Marschall, 2022), we have the following assumption: individuals may (a) be motivated to participate in an election—a process known as mobilisation, (b) have their existing voting decision reinforced, or (c) alter their intended vote choice, a phenomenon referred to as conversion. We consider this heuristic equally applicable to learning effects resulting from the use of the Wahl-O-Mat, as this VAA is widely used in election campaigns. Again, van de Pol et al. (2014) highlight how VAA usage is often perceived as a confirmation of existing preferences, functioning as a form of checking rather than changing voter behaviour. Consequently, we assume that a reinforcement effect is also relevant in this context. Mobilisation, reinforcement, and conversion are thus key dimensions of political behaviour that may be influenced by the use of VAA's. This debate leads us to our first hypothesis:

H1: There are significant differences between individuals with lower levels of formal education and those with higher levels of formal education in reporting that their political behaviour is influenced by VAA usage.

Given the role of prior political information, we assume that the interaction between VAA usage and political learning could mediate the relationship between education and political behaviour. Less educated voters, who often exhibit lower levels of political interest and efficacy, may particularly benefit from the educational potential of VAA's. Mahéo (2017, p. 511) highlights this dynamic, arguing that “VAA's work differently for more or less educated citizens,” with the interaction between attention and educational level showing a positive relationship—yet only for less educated individuals in their sample. Building on this, we expect that the potential of VAA's to enhance political learning is particularly pronounced among less educated individuals. Consequently, this learning effect may act as a mediating factor, shaping their political behaviour:

H2: Individuals with lower levels of formal education are more likely to experience a learning effect from VAA usage, which in turn mediates the influence of VAA usage on their political behaviour compared to individuals with higher levels of formal education.

3. Research Design

3.1. Case Selection

The Wahl-O-Mat has become one of the most established VAAs and is now an almost indispensable feature of the pre-election public sphere—before the 2021 German Federal Election, it was used over 21.3 million times. Since its introduction in 2002, the Wahl-O-Mat has steadily gained popularity, achieving a high level of recognition and significant political relevance in Germany (Albertsen, 2022; Wurthmann & Marschall, 2022).

At the same time, the 2021 German Federal Election differed from previous ones in its exceptional openness. It remained uncertain which party would ultimately emerge as the frontrunner, who would assume the Chancellery, and which coalition would form the next government. This uncertainty heightened public interest in how voters, particularly amid the ongoing Covid-19 pandemic, sought political information and ultimately arrived at their electoral decisions (Wurthmann & Marschall, 2022).

Against the backdrop of increasing party system heterogeneity, growing fragmentation, and concurrent uncertainty surrounding political processes in the Federal Republic of Germany (Debus, 2022), VAAs like the Wahl-O-Mat serve as crucial mediators (Garzia & Marschall, 2019). Given this context, it is worth examining the role such a widely used tool plays—especially in times of political and social uncertainty.

3.2. Data

The data utilised in this study were collected as part of a four-wave panel survey on the 2021 German Federal Election, designed to measure public attitudes and political behaviour across multiple waves. The survey was conducted using the Bilendi online access panel, with respondent recruitment based on a quota sampling method aligned with population-representative parameters.

The first wave of the survey was strategically conducted between 9 and 14 August, prior to the final phase of the campaign and voter mobilisation efforts. A total of 2,270 respondents participated, of whom 2,206 were eligible voters. The second wave, administered between 17 and 25 September—just one day before the Bundestag election—secured responses from 1,533 individuals, reflecting a panel mortality rate of about 30.5 percent.

As outlined in Section 2, this article seeks to examine the underlying mechanisms through which the Wahl-O-Mat influences voter behaviour, whether by mobilising individuals to participate in the election, reinforcing existing voting intentions, or prompting a switch from one party to another. Given this focus, familiarity with the Wahl-O-Mat is a prerequisite for inclusion in the analysis. Among the second-wave respondents, 1,235 respondents met this criterion, necessitating the exclusion of 298 individuals who were unfamiliar with the tool. Furthermore, 976 respondents had actively used the Wahl-O-Mat at some point, leading to the exclusion of an additional 259 who had not.

Among the remaining 976 respondents, 260 were excluded as they had previously used the Wahl-O-Mat, but not specifically for the 2021 Federal Election. This left a sample of 716 cases before the final data cleaning process. At this stage, respondents who failed to answer any of the key indicators employed in the study or who provided only non-substantive responses were further excluded. Ultimately, this refinement resulted in a final analytical sample of 667 respondents.

Regarding the indicators used in the analysis, demographic variables such as age, formal education level, sex, and place of residence were drawn from Survey Wave 1. In contrast, variables susceptible to campaign influences, including party identification and political interest, were derived from Survey Wave 2. The same applies to Wahl-O-Mat-related factors, which were only introduced in the second wave, as the tool was launched online after the initial survey wave had been completed.

3.3. Operationalisation

3.3.1. Dependent Variables

We have introduced our threefold concept of political behaviour, which we assume is influenced by the use of VAAs. We operationalise our indicators as follows:

- (1) Mobilisation: The Wahl-O-Mat motivated me to participate in the federal election.
- (2) Reinforcement: The Wahl-O-Mat reinforced my existing voting decision.
- (3) Conversion: As a result of the Wahl-O-Mat outcome, I am likely to change my vote choice.

Respondents were asked to express their position using the scale (1) *strongly disagree*, (2) *disagree*, (3) *neither agree nor disagree*, (4) *agree*, and (5) *strongly agree*.

3.3.2. Independent and Mediator Variables

The central variable of interest in this case is the formal educational level of the respondents. Respondents were presented with a series of options to classify their respective educational qualifications. They were provided with the following selection:

- (1) Finished school without school-leaving certificate.
- (2) Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling ("Hauptschulabschluss," "Volksschulabschluss").
- (3) Intermediary secondary qualification, after 10 years of schooling ("Mittlere Reife," "Realschulabschluss"/"Polytechnische Oberschule mit Abschluss 10. Klasse").
- (4) Certificate fulfilling entrance requirements to study at a polytechnic ("Fachhochschulreife [Abschluss einer Fachoberschule etc.]").
- (5) Higher qualification, entitling holders to study at a university ("Abitur"/"Erweiterte Oberschule mit Abschluss 12. Klasse [Hochschulreife]").
- (6) Another type of school leaving qualification, specifically: _____ (free text field).
- (9) I am still attending school.

Responses provided under option (6) were recoded within the existing coding scheme whenever possible. If recoding was not feasible, these responses were omitted from the analysis. Additionally, since option (9) does not provide information on previously obtained qualifications, respondents selecting this option were excluded from further analysis. To create a dichotomous variable for educational attainment, responses from options (2) and (3) were grouped into the category *No higher education* (coded as 1), while responses from options (4) and (5) were classified as *Higher education* (coded as 0). Collapsing options (2) and (3) was necessary because, after panel attrition and data cleaning, only 55 respondents remained in category (3) compared to 417 in category (1) and 240 in category (2). A separate handling of options (2) and (3) would not have allowed for reliable analyses, particularly given the complexity of the path modelling approach employed. As a result of panel attrition and data cleaning, no respondents who had completed schooling without obtaining a diploma remained in the final dataset.

Additionally, we incorporate two mediator variables, which we expect to be influenced by formal educational attainment and, in turn, affect the dependent variables. These mediators capture the learning effects that respondents report when using the Wahl-O-Mat.

The first mediator (awareness) reflects whether respondents became aware of party positions that were previously unknown to them. The second (party differentiation) measures respondents' perceived ability to discern differences between parties as a result of using the Wahl-O-Mat. The indicators employed for these variables are as follows:

- (1) Party Position Awareness: The Wahl-O-Mat made me aware of party positions that I was previously unaware of.
- (2) Party Differences: The Wahl-O-Mat helped me identify differences between parties.

Participants were requested to indicate their stance on a scale ranging: (1) *strongly disagree*, (2) *disagree*, (3) *neither agree nor disagree*, (4) *agree*, and (5) *strongly agree*.

While awareness of party positions and awareness of party differences are conceptually connected—since recognising differences arguably presupposes at least some knowledge of positions—we treat them analytically as distinct mediators. This distinction enables us to assess whether both aspects of perceived learning exert independent effects on political behaviour. A strictly sequential relationship in which position awareness precedes difference awareness would indeed reflect a plausible cognitive process. However, explicitly modelling this sequence would substantially increase the complexity of the analytical framework and introduce potential endogeneity with other covariates (e.g., party identification). We therefore maintain the specification presented here, while noting that the two mediators are moderately correlated (see Figure A5 in the Supplementary File).

3.3.3. Controls

First, we address key variables by considering respondents' age, sex, and place of residence, which have been consistently shown to be highly relevant for political behaviour (Cramer Walsh, 2012; Smets & van Ham, 2013). Additionally, we control for the effects of party identification and political interest, as both have been demonstrated to significantly influence the use of the Wahl-O-Mat and its subsequent implications (Wurthmann & Marschall, 2022).

Regarding sex, we distinguish between (0) male and (1) female. Age was recorded in an open-ended format, with respondents reporting values ranging from 18 to 74 years. For place of residence, respondents selected from a five-point scale: (1) large city, (2) outskirts or suburbs of a large city, (3) medium or small town, (4) rural village, and (5) isolated house in the countryside. Due to the small number of cases in category (5), it has been merged with category (4).

For party identification, respondents were asked whether they identify themselves with a political party and, if so, which one. The variable is coded as (1) for all respondents who indicated a party affiliation and (0) for those who reported no party identification. Political interest is measured on a five-point scale, ranging from (1) *not interested in politics at all* to (5) *very strongly interested in politics*.

3.4. Method and Analysis Strategy

We proceed as follows. First, we present the descriptive findings by examining the distributions of the dependent and mediator variables for the entire sample. Subsequently, we differentiate between the two educational groups based on our dichotomous classification: (0) higher education and (1) no higher education.

Next, we conduct bivariate analyses to explore the relationships between the independent variable (education) and both the dependent and mediator variables. This step provides an initial indication of potential associations before proceeding to more complex modelling.

To further investigate these relationships, we employ path models within the framework of Structural Equation Modelling. This approach enables us to decompose direct and indirect effects, offering a more nuanced understanding of how education relates to the dependent variables. Specifically, we hypothesize that the relationship between formal educational attainment and political behaviour is partially mediated by the learning effects attributed to the Wahl-O-Mat—namely, awareness (whether respondents discovered previously unknown party positions) and differentiation (their ability to discern differences between parties).

Path models are particularly well-suited for this analysis for several reasons. First, they allow us to distinguish direct effects (education's influence on the dependent variables) from indirect effects (mediated by awareness and differentiation). Second, they enable us to test specific mediation mechanisms, assessing whether Wahl-O-Mat learning effects play a role in shaping the influence of education or whether its impact remains largely direct. Third, this method accounts for reciprocal interdependencies between mediators and outcomes, ensuring that observed relationships are not oversimplified.

By integrating descriptive, bivariate, and path model analyses, we establish a comprehensive framework for understanding how differences concerning educational attainment relate to political behaviour via cognitive mechanisms associated with Wahl-O-Mat use. This approach provides deeper insights than standard regression models, which may overlook these underlying patterns of association.

4. Results and Discussion

Given that the Wahl-O-Mat is predominantly used by individuals already possessing higher levels of political interest compared to the general population (Albertsen, 2022; Marschall, 2011), it is unsurprising that the

tool is not widely perceived as a mobilising force. Among respondents, 46.5 percent strongly disagree and 21.0 percent disagree with the statement that the Wahl-O-Mat motivated them to participate in the federal election. An additional 13.6 percent express neutrality, while only 11.4 percent agree and 7.5 percent strongly agree. However, the perception shifts significantly when considering the reinforcement of voting decisions. Here, only 17.8 percent strongly disagree and 11.1 percent disagree that the Wahl-O-Mat influenced their decision. By contrast, 21.9 percent remain neutral, while 25.9 percent agree and 23.2 percent strongly agree, suggesting a notable role of the tool in strengthening pre-existing vote choice. When it comes to conversion—where respondents indicate adjusting their vote based on the Wahl-O-Mat's outcome—the pattern resembles that of mobilisation but with even stronger opposing views. A substantial 57.9 percent strongly disagree and 21.1 percent disagree with the idea that the tool changed their vote choice. Additionally, 11.8 percent are neutral, while only 6.3 percent agree and a mere 2.8 percent strongly agree (see Figure A1 in the Supplementary File). Overall, these findings suggest that while the Wahl-O-Mat is not perceived as a catalyst for changing vote choice, it serves primarily as a tool for reinforcing pre-existing vote choices.

Furthermore, it is worth examining how respondents perceive the self-attributed effects of using the Wahl-O-Mat. Regarding awareness of party positions, 24.7 percent strongly disagree, and 16.3 percent disagree with the statement that the tool made them aware of previously unknown positions. An additional 23.0 percent remain neutral, while just over 27.0 percent agree, and nearly 9.0 percent strongly agree. A different pattern emerges when considering the perceived ability to recognise differences between party positions. Here, significantly fewer respondents reject the notion that the Wahl-O-Mat enhanced their understanding. Only 15.4 percent strongly disagree and 15.1 percent disagree with the statement. Meanwhile, 30.9 percent express neutrality, while 29.5 percent agree and 9.0 percent strongly agree (see Figure A2 in the Supplementary File).

While a substantial share of respondents does not attribute new issue awareness to the Wahl-O-Mat, many acknowledge its role in highlighting differences between parties, underlining its function as a comparative tool rather than a source of entirely new political information.

The frequency distributions, segmented by individuals with and without formal higher education levels, reveal no substantial differences. The same holds true for the average effects attributed to the Wahl-O-Mat. While the two groups exhibit statistically significant differences ($p < 0.001$), the magnitude of these effect differences remains relatively modest (see Table 1).

Table 1. Mean values comparison.

	High education	Low/medium education
Mobilisation	2.12 (95% c.i. = 1.99, 2.25)	2.13 (95% c.i. = 1.97, 2.29)
Reinforcement	3.32 (95% c.i. = 3.18, 3.45)	3.17 (95% c.i. = 3.00, 3.34)
Conversion	1.71 (95% c.i. = 1.61, 1.82)	1.80 (95% c.i. = 1.67, 1.94)
Party position awareness	2.81 (95% c.i. = 2.68, 2.94)	2.77 (95% c.i. = 2.61, 2.92)
Party position differences	3.01 (95% c.i. = 2.89, 3.12)	3.03 (95% c.i. = 2.88, 3.17)

Note: $N = 667$.

This impression is further reinforced by a subsequent correlational analysis. While all mediator and dependent variables exhibit highly significant correlations with one another ($p < 0.001$), no significant effect is observed between education and any of the respective determinants (see Figure A5 in the Supplementary File).

At first glance, this finding may not be surprising, particularly in light of Albertsen (2022, p. 407), who identified a diminishing effect of higher education levels on Wahl-O-Mat usage. However, it raises the question of how these relationships unfold when accounting for the previously mentioned control variables.

Our data analysis aims to examine both direct and indirect effects on three key aspects of electoral behaviour: mobilisation, reinforcement, and conversion. Using path models, our analysis demonstrates that individuals with lower levels of education are significantly more likely to report that the Wahl-O-Mat increased their awareness of differences between parties and their positions. This heightened awareness, in turn, leads to a greater sense of mobilisation. As expected, no direct effect of education on mobilisation is observed. These findings reinforce the idea that the Wahl-O-Mat enhances perceived competence by facilitating incremental knowledge acquisition, which ultimately contributes to a higher perceived level of mobilisation. Notably, the total effect of education, when mediated by awareness of party differences, is slightly stronger (0.129 **; see Figure 1) compared to the effect mediated by awareness of party positions more broadly (0.109 *; see Figure 1).

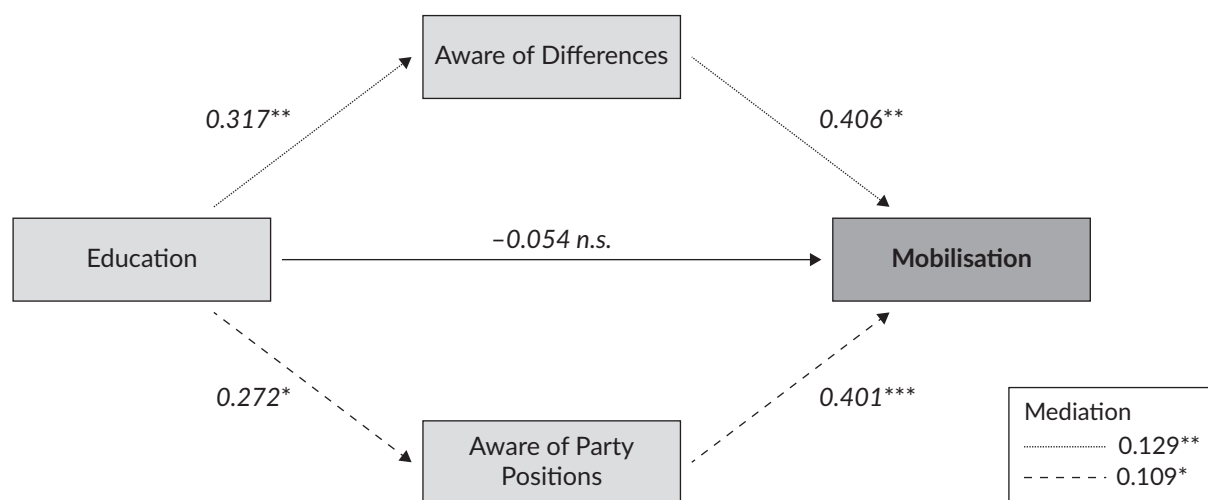


Figure 1. Mediated effects on mobilisation. Notes: $N = 667$ (see Table A1 in the Supplementary File for the full results), + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Similar patterns emerge when examining reinforcement. Individuals with lower or middle levels of formal education are more likely than those with higher levels to perceive a competence gain through the Wahl-O-Mat, which in turn strengthens their sense of reinforcement. However, an interesting contrast arises: A negative direct effect is observed between lower education levels and reinforcement. How can this be explained? One plausible reason is that individuals with higher levels of formal education tend to have stronger pre-existing voting dispositions, as previously discussed. Moreover, research on VAAs, such as the study by van de Pol et al. (2014), suggests that highly educated individuals not only report lower learning effects but also primarily use these tools for confirmation rather than for knowledge gain. This group of so-called “checkers”—who engage with VAAs to validate rather than reassess their preferences—consists predominantly of highly educated individuals, accounting for well over two-thirds. Given our descriptive

findings, this result is not surprising. However, it does challenge our initial assumption that direct effects would be negligible (see Figure 2). Notably, the total effect of education, when mediated by awareness of party differences, is considerably stronger (0.147 **; see Figure 2) compared to the effect mediated by awareness of party positions more broadly (0.068 *; see Figure 2).

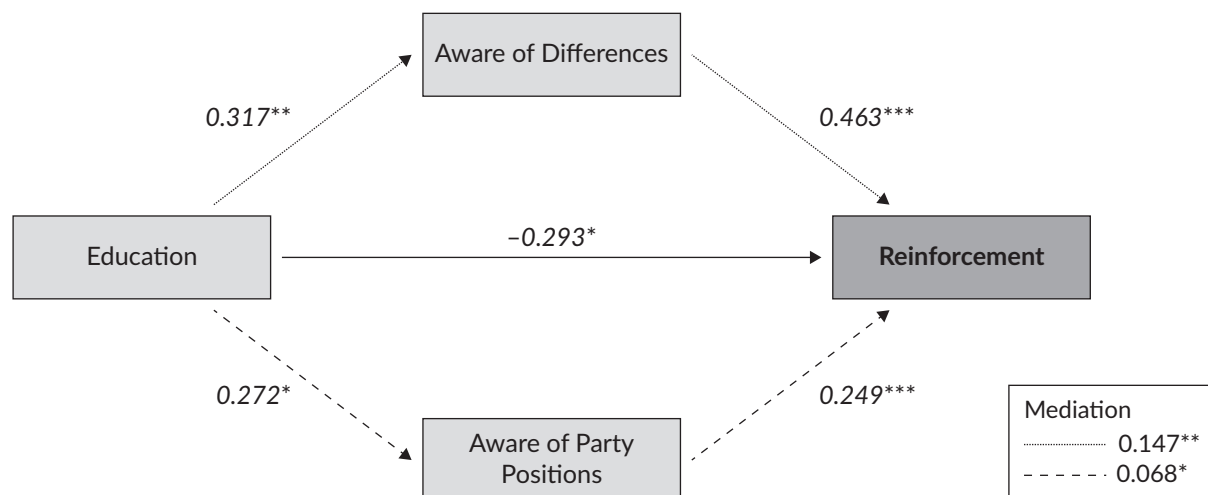


Figure 2. Mediated effects on reinforcement. Notes: $N = 667$ (see Table A2 in the Supplementary File for the full results), + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Regarding self-reported conversion induced by the Wahl-O-Mat, the effects closely resemble those observed for mobilisation. Here, too, indirect effects are clearly present, while no direct effects are found. Notably, among the indirect effects, the impact mediated by awareness of differences between parties (0.085 **; see Figure 3) is consistently stronger across all models compared to the effect mediated by learning about party positions per se (0.071 *; see Figure 3).

Individuals with lower levels of formal education do not measurably report being mobilised by the Wahl-O-Mat, being influenced in their party choice, or feeling reinforced in their voting decisions more than

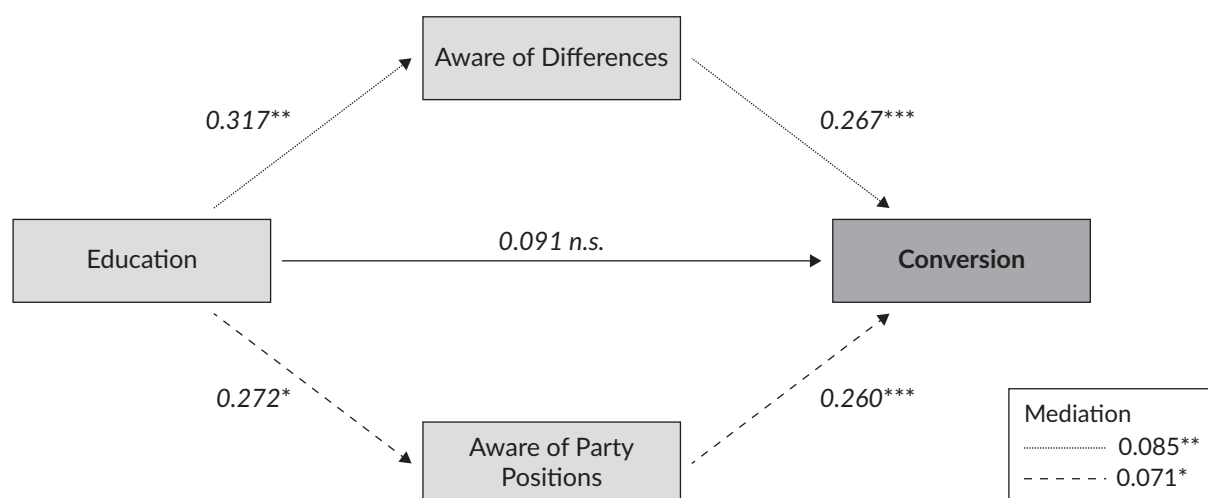


Figure 3. Mediated effects on conversion. Notes: $N = 667$ (see Table A3 in the Supplementary File for the full results), + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

those with higher education levels. Nor do they attribute a general impact on their opinion-forming to the tool. However, individuals with lower levels of education are significantly more likely than those with higher levels to report learning effects from the Wahl-O-Mat. This perceived knowledge gain is strongly associated with a greater sense of mobilisation, reinforcement, conversion, and general influence attributed to the tool. Consequently, when individuals with lower levels of education experience these learning effects, they assign greater importance to the Wahl-O-Mat overall. These findings highlight the varying roles of educational background in shaping how individuals engage with political information and educational tools such as the Wahl-O-Mat.

Regarding H1, which posited that individuals with lower levels of formal education differ significantly from those with higher levels, our results present a mixed picture. While no significant differences emerge between the education groups in terms of mobilisation or conversion, notable variations are observed for reinforcement and the general influence attributed to the Wahl-O-Mat. A negative direct effect is observed for highly educated individuals regarding reinforcement, which may be explained by the tendency of highly educated users to use the Wahl-O-Mat primarily to confirm their existing voting preferences—a behaviour known as “checking” (van de Pol et al., 2014). In contrast, highly educated individuals report a significantly greater general influence of the tool on their behaviour than their less-educated counterparts. However, the mediated effects provide a different conclusion, suggesting that indirect mechanisms play an important role.

This brings us to H2, which hypothesised that individuals with lower levels of education are more likely to experience learning effects from VAA usage, which in turn mediates their political behaviour. Our findings strongly confirm this hypothesis. The data consistently indicate that individuals with lower levels of education experience more pronounced learning effects when using the Wahl-O-Mat. These learning effects, in turn, significantly mediate the relationship between VAA usage and changes in political behaviour, demonstrating that VAAs can play a crucial role in enhancing political knowledge and influencing political engagement among these users.

Overall, Germany—with its former division into West and East—might exhibit a regional bias regarding the effects described above. Distinct regional political cultures, varying levels of institutional trust, and diverging patterns of political engagement and voting behaviour (Pickel & Pickel, 2023; Walke & Wurthmann, 2024) may influence the use of and trust in VAAs such as the Wahl-O-Mat. However, when including place of residence—coded as Western Germany (0) and Eastern Germany (1)—our results remain robust and largely unchanged (see Tables A4–A6 in the Supplementary File).

5. Conclusion

In conclusion, this study illuminates the nuanced impact of VAAs on individuals with varying levels of formal education. Given the importance of education for political interest, knowledge, and engagement, the study’s objective was to examine this key factor in the usage of VAAs. While one might assume that VAAs like the Wahl-O-Mat are predominantly used by highly educated individuals and therefore, pointedly put, a closer look at less educated users seems to be less instructive, empirical findings indicate (a) that their impact extends across different educational backgrounds, influencing both highly and less educated users, and (b) that educational level makes a difference when it comes to mechanism of the effect of VAA usage. The findings reveal that while individuals with higher education levels are more likely to perceive a general

influence of VAAs on their voting behaviour, those with lower or middle education levels report greater learning effects from using such tools. This learning experience, in turn, leads to a perceived increase in competence and a higher likelihood of feeling mobilised, reinforced, or politically converted by the VAA.

The study has several shortcomings. One notable limitation is that the analysis of the VAA users was conducted before a very special election held during the pandemic, which may have generated influences not to be found in typical election cycles. Additionally, the study relies on self-reporting, which can introduce biases and inaccuracies, as participants' responses may be influenced by their perceptions, memory, and willingness to disclose accurate information. This also raises a more substantial issue of measurement validity. All key outcomes of this study—including learning effects, mobilisation, and vote change—are based on self-reporting. While such indicators are valuable for capturing participants' subjective experiences, they are vulnerable to social desirability and recall biases. Without objective measures of political knowledge or independent validation of actual behaviour, it remains uncertain whether the reported effects reflect real changes or primarily respondents' self-assessments. This is obviously not a VAA-related measurement problem but rather a fundamental question of measurement methodology in the field of effect research. Another significant shortcoming of the study might be the question of its generalisability to other national contexts. Germany has a unique party system, electoral system, and "VAA culture." This context is shaped by comparatively high levels of digital literacy and a comparatively stable multiparty system, both of which may increase the salience and accessibility of VAAs. Therefore, the question remains open as to whether the findings from this study, conducted in such a distinctive environment, can be effectively applied to other countries with different political landscapes, electoral systems, and levels of VAA integration. Moreover, the timing of the study poses an additional limitation. The 2021 German Federal Election was held during the Covid-19 pandemic, a period marked by elevated political uncertainty, widespread reliance on digital tools, and unusual campaign dynamics. It is possible that these extraordinary circumstances modified the relevance and use of VAAs in ways that may not recur in typical electoral cycles (Wurthmann & Marschall, 2022). These issues call for more comparative research. Comparative designs could provide evidence on how institutional and cultural settings condition the effects of VAAs, while longitudinal approaches could test whether the patterns observed here are valid beyond the specific context of the 2021 German Federal Election. By focusing on the educational aspects and their mediated effects, VAA research can more effectively contribute to measuring the contribution of these tools, also for the democratic process. Future research should continue to explore the varying effects of VAAs across various demographic segments.

Another limitation concerns sampling and sample attrition. The use of an online access panel enables us to reach a broader spectrum of respondents compared to widespread university-based samples, but this approach does not fully resolve the problem of representativeness. As there are currently no systematic data on the socio-demographic profile of Wahl-O-Mat users, we cannot check to what extent the sample reflects the broader user base. Potential self-selection into the panel between survey waves may bias the composition of respondents toward individuals who are politically more interested and who possess higher levels of formal education.

Moreover, although the study employs a robust four-wave panel design, the final sample was reduced to 667 respondents. In particular, panel mortality—as in our study—is typically correlated with political interest and formal education (Müller, 2025). Several respondents had to be excluded because they either did not know the Wahl-O-Mat or had never used the tool before. Such patterns, however, are also strongly shaped

by levels of political interest, knowledge, and education, which are themselves closely interrelated. This raises the possibility of attrition bias, particularly along educational lines, which are central to our analysis. Prior research has noted that VAAs are disproportionately used by better-educated citizens who often employ them to confirm existing preferences rather than to reconsider them (van de Pol et al., 2014).

The dynamics described may have reinforced the already existing educational skew. This is also reflected in the fact that we could not differentiate more finely within the lower education categories and had to merge them into a single group to provide analytical reliability. Future research should seek to differentiate more finely between different levels of education, as this would allow a deeper understanding of how VAAs operate across more differentiated educational backgrounds. To sum it up: A more finely grained operationalisation of education could reveal additional variation in political behaviour.

Finally, there is some significant learning that makers of these tools can draw from the study. The findings underscore the importance of designing VAA research to consider its educational components in detail, particularly for users with lower educational backgrounds. As the learning experience provided by tools such as the Wahl-O-Mat seems to be particularly important for individuals with medium or low levels of formal education, the specific design of VAAs should take didactical approaches more strongly into consideration. By doing so, VAAs might be much more attractive and capable of not only “preaching to the converted” but also of reaching out to those who are much more in need of political information and orientation. Indeed, we started with the assumption that people with low levels of education would experience only a low impact of using VAAs, but our study indicates that this is not the case; rather, the opposite.

Acknowledgments

The authors gratefully acknowledge the financial support for publication provided by the University and State Library (ULB) of Heinrich Heine University Düsseldorf. In this article, editorial language assistance was provided by Don Mac Donald (c/o English Language Work, Dartmouth, Nova Scotia, Canada B2Y 1M2).

Funding

The data used for this article were collected in a project of the Chair of Political Science II at the Heinrich Heine University Düsseldorf. The authors wish to acknowledge the funding of the data collection by the Fritz Thyssen Foundation (Az. 20.23.0.007PO).

Conflict of Interests

The authors declare no conflict of interests. In this article, editorial decisions were undertaken by Diego Garzia (University of Lausanne / Bologna University), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University).

Data Availability

Research data can be requested via the project link and the responsible persons listed there: <https://www.sozwiss.hhu.de/institut/abteilungen/politikwissenschaft/politik-ii/prof-dr-stefan-marschall/wahlen-in-stuermischen-zeiten>

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Abou-Chadi, T., & Hix, S. (2021). Brahmin left versus merchant right? Education, class, multiparty competition, and redistribution in Western Europe. *The British Journal of Sociology*, 72(1), 79–92. <https://doi.org/10.1111/1468-4446.12834>
- Albertsen, A. (2022). How do the characteristics of voting advice application users change over time? Evidence from the German election studies. *German Politics*, 31(3), 399–419. <https://doi.org/10.1080/09644008.2020.1830374>
- Bornschiefer, S., Haffert, L., Häusermann, S., Steenbergen, M. R., & Zollinger, D. (2024). *Cleavage formation in the twenty-first century*. Cambridge University Press. <https://doi.org/10.5167/uzh-265992>
- Bourdieu, P., & Passeron, J. C. (1990). *Reproduction in education, society and culture* (Vol. 4). Sage.
- Breen, R., Luijckx, R., Müller, W., & Pollak, R. (2009). Nonpersistent inequality in educational attainment: Evidence from eight European countries. *American Journal of Sociology*, 114(5), 1475–1521. <https://doi.org/10.1086/595951>
- Cramer Walsh, K. (2012). Putting inequality in its place: Rural consciousness and the power of perspective. *American Political Science Review*, 106(3), 517–532. <https://doi.org/10.1017/S0003055412000305>
- Dalton, R. J. (2012). *The apartisan American: Dealignment and the transformation of electoral politics*. CQ Press.
- Dassonneville, R., & Hooghe, M. (2017). Voter turnout decline and stratification: Quasi-experimental and comparative evidence of a growing educational gap. *Politics*, 37(2), 184–200. <https://doi.org/10.1177/0263395716674370>
- Debus, M. (2022). Parteienwettbewerb und Wahrscheinlichkeit verschiedener Koalitionsoptionen bei der Bundestagswahl 2021. *Politische Vierteljahresschrift*, 63(1), 73–88. <https://doi.org/10.1007/s11615-021-00361-8>
- de Jong, J., & Kamphorst, J. (2025). Separated by degrees: Social closure by education levels strengthens contemporary political divides. *Comparative Political Studies*, 58(7), 1533–1568. <https://doi.org/10.1177/00104140241271104>
- Dieing, T. I. (2025). Just can't get enough—Profiling users of multiple voting advice applications. *Journal of Information Technology & Politics*. Advance online publication. <https://doi.org/10.1080/19331681.2025.2496259>
- Elff, M., & Roßteutscher, S. (2022, June 23–25). *Education as a new cleavage? The relevance of class, education, and income in a long-term perspective* [Paper presentation]. EPSA Conference 2022, Prague, Czech Republic.
- Enyedi, Z. (2015). The influence of voting advice applications on preferences, loyalties and turnout: An experimental study. *Political Studies*, 64(4), 1000–1015. <https://doi.org/10.1111/1467-9248.12213>
- Garzia, D. (2010). The effects of VAAs on users' voting behaviour: An overview. In L. Cedroni & D. Garzia (Eds.), *Voting advice applications in Europe: The state of the art* (pp. 13–33). Scriptaweb.
- Garzia, D., & Marschall, S. (2016). Research on voting advice applications: State of the art and future directions. *Policy & Internet*, 8, 376–390. <https://doi.org/10.1002/poi3.140>
- Garzia, D., & Marschall, S. (2019). Voting advice applications. In W. R. Thompson (Eds.), *Oxford research encyclopedia of politics* (pp. 1–22). Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228637.013.620>
- 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., & Gemenis, K. (2023). Do voting advice applications affect party preferences?

- Evidence from field experiments in five European countries. *Political Communication*, 40(3), 597–614. <https://doi.org/10.1080/10584609.2023.2181896>
- Hansen, K. M., & Pedersen, R. T. (2014). Campaigns matter: How voters become knowledgeable and efficacious during election campaigns. *Political Communication*, 31(2), 303–324. <https://doi.org/10.1080/10584609.2013.815296>
- Häusermann, S., & Kriesi, H. (2015). What do voters want? Dimensions and configurations in individual-level preferences and party choice. In P. Beramendi, S. Häusermann, H. Hitschelt, & H. Kriesi (Eds.), *The politics of advanced capitalism* (pp. 202–230). Cambridge University Press.
- Heinsohn, T., Israel, J., Marschall, S., & Schultze, M. (2016). Online-Wahlhilfen in Wahlkämpfen. Empirische Ergebnisse einer Panelbefragung zur Europawahl 2014. *Zeitschrift für Politikwissenschaft*, 26, 253–277. <https://doi.org/10.1007/s41358-016-0055-y>
- Hooghe, L., & Marks, G. (2025). How does the education cleavage stack up against the classic cleavages of the past? *West European Politics*. Advance online publication. <https://doi.org/10.1080/01402382.2025.2452789>
- Israel, J., Marschall, S., & Schultze, M. (2017). Cognitive dissonance and the effects of voting advice applications on voting behaviour: Evidence from the European elections 2014. *Journal of Elections, Public Opinion and Parties*, 27(1), 56–74. <https://doi.org/10.1080/17457289.2016.1268142>
- Kamoen, N., Holleman, B., Krouwel, A., Van de Pol, J., & De Vreese, C. (2017). The effect of voting advice applications on political knowledge and vote choice. In F. Buckley & T. Reidy (Eds.), *Electoral management: Institutions and practices in an established democracy* (pp. 161–184). Routledge.
- Koch, C. M., Meléndez, C., & Rovira Kaltwasser, C. (2021). Mainstream voters, non-voters and populist voters: What sets them apart? *Political Studies*, 71(3), 893–913. <https://doi.org/10.1177/00323217211049298>
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice: How the voter makes up his mind in a presidential campaign*. Columbia University Press.
- Mahéo, V.-A. (2017). Information campaigns and (under)privileged citizens: An experiment on the differential effects of a voting advice application. *Political Communication*, 34, 511–529. <https://doi.org/10.1080/10584609.2017.1282560>
- Manavopoulos, V., Triga, V., Marschall, S., & Wurthmann, L. C. (2018). The impact of VAAs on (non-voting) aspects of political participation: Insights from panel data collected during the 2017 German federal elections campaign. *Statistics, Politics and Policy*, 9(2), 105–134. <https://doi.org/10.1515/spp-2018-0008>
- Marks, G., Attewell, D., Hooghe, L., Rovny, J., & Steenbergen, M. (2023). The social bases of political parties: A new measure and survey. *British Journal of Political Science*, 53(1), 249–260. <https://doi.org/10.1017/S0007123421000740>
- Marschall, S. (2011). Nutzer und Nutzen—Der Wahl-O-Mat zur Bundestagswahl 2009. In E. J. Schweitzer & S. Albrecht (Eds.), *Das Internet im Wahlkampf: Analysen zur Bundestagswahl 2009* (pp. 136–153). Springer. https://doi.org/10.1007/978-3-531-92853-1_5
- Marschall, S. (2014). Profiling users. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates: Voting advice applications in comparative perspective* (pp. 93–104). ECPR Press.
- Marschall, S., & Israel, J. (2014). Toy or tool? Der Wahl-O-Mat als hybrides Angebot der politischen Bildung. *GWP-Gesellschaft. Wirtschaft. Politik*, 63(3), 365–378.
- Marschall, S., & Schultze, M. (2012). Normalisierung oder Mobilisierung?—Die Auswirkungen politischer Online-Kommunikation auf die Wahlbeteiligung am Beispiel einer Internet-Applikation zur Bundestagswahl 2009. *Politische Vierteljahresschrift*, 53(3), 444–466.
- Mayer, S. J., & Schultze, M. (2019). The effects of political involvement and cross-pressures on multiple party

- identifications in multi-party systems—Evidence from Germany. *Journal of Elections, Public Opinion and Parties*, 29(2), 245–261. <https://doi.org/10.1080/17457289.2018.1466785>
- 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>
- Müller, K. (2025). Survey nonresponse after elections: Investigating the role of winner-loser effects in panel attrition. *International Journal of Public Opinion Research*, 37(3), Article edaf031.
- Munzert, S., Barberá, P., Guess, A. M., & Yang, J. (2020). Do online voter guides empower citizens? Evidence from a field experiment with digital trace data. *Public Opinion Quarterly*, 84(3), 675–698. <https://doi.org/10.1093/poq/nfaa037>
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706. <https://doi.org/10.1080/10584609.2020.1843572>
- Nie, N. H., Verba, S., & Petrocik, J. R. (1979). *The changing American voter*. Harvard University Press.
- Pfeffer, F. T. (2008). Persistent inequality in educational attainment and its institutional context. *European Sociological Review*, 24(5), 543–565. <https://doi.org/10.1093/esr/jcn026>
- Pickel, S., & Pickel, G. (2023). The wall in the mind—Revisited stable differences in the political cultures of Western and Eastern Germany. *German Politics*, 32(1), 20–42. <https://doi.org/10.1080/09644008.2022.2072488>
- Plutzer, E. (2002). Becoming a habitual voter: Inertia, resources, and growth in young adulthood. *American Political Science Review*, 96(1), 41–56. <https://doi.org/10.1017/S0003055402004227>
- Schäfer, A., Roßteutscher, S., & Abendschön, S. (2020). Rising start-up costs of voting: Political inequality among first-time voters. *West European Politics*, 43(4), 819–844. <https://doi.org/10.1080/01402382.2019.1638141>
- Schäfer, A., & Steiner, N. D. (2025). How education, generation and gender jointly structure green and radical right voting. *West European Politics*. Advance online publication. <https://doi.org/10.1080/01402382.2025.2466122>
- Scott, R. (2022). Does university make you more liberal? Estimating the within-individual effects of higher education on political values. *Electoral Studies*, 77, Article 102471. <https://doi.org/10.1016/j.electstud.2022.102471>
- Smets, K., & van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral Studies*, 32(2), 344–359. <https://doi.org/10.1016/j.electstud.2012.12.006>
- Stadelmann-Steffen, I., Rajski, H., & Ruprecht, S. (2023). The role of vote advice applications in direct-democratic opinion formation: An experiment from Switzerland. *Acta Politica*, 58(4), 792–818. <https://doi.org/10.1057/s41269-022-00264-5>
- Stubager, R. (2013). The development of the education cleavage. Denmark as a critical case. In Z. Enyedi & K. Deegan-Krause (Eds.), *The structure of political competition in Western Europe* (pp. 91–119). Routledge.
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15, 582–599. <https://doi.org/10.1017/S1755773923000103>
- van de Pol, J. (2016). *Voting wiser: The effect of voting advice applications on political understanding*. University of Amsterdam.
- 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 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>

Verba, S., Scholzman, K. L., & Brady, H. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Walke, L., & Wurthmann, L. C. (2024). Koalitionen mit der AfD? Determinanten der Präferenzbildung in Ost- und Westdeutschland vor der Bundestagswahl 2021. *Zeitschrift für Vergleichende Politikwissenschaft*, 18(3), 401–426. <https://doi.org/10.1007/s12286-024-00619-6>

Wall, M., Krouwel, A., & Vitiello, T. (2014). Do voters follow the recommendations of voter advice application websites? A study of the effects of kieskompas.nl on its users' vote choices in the 2010 Dutch legislative elections. *Party Politics*, 20(3), 416–428. <https://doi.org/10.1177/1354068811436054>

Westheuser, L., & Zollinger, D. (2025). Cleavage theory meets Bourdieu: Studying the role of group identities in cleavage formation. *European Political Science Review*, 17(1), 110–127. <https://doi.org/10.1017/S1755773924000249>

Wurthmann, L. C., & Marschall, S. (2022). Kommunikation und Komplexität politischer Informationssuche in der Pandemie: Der Wahl-O-Mat, Social Media und klassische Informationsressourcen im Bundestagswahljahr 2021. In K.-R. Korte, M. Schifers, A. von Schuckmann, & S. Plümer (Eds.), *Die Bundestagswahl 2021: Analysen der Wahl-, Parteien-, Kommunikations- und Regierungsforschung* (pp. 1–26). Springer. https://doi.org/10.1007/978-3-658-35758-0_4-1

About the Authors



L. Constantin Wurthmann is a research fellow at the Mannheim Centre for European Social Research (MZES), University of Mannheim. His research, focusing on electoral behaviour, party competition, and political representation, has been published in journals such as the *European Political Science Review*, *Electoral Studies*, *Party Politics*, *JEPOP*, and *Political Research Exchange*.



Daniel C. Hagemann is a research associate and doctoral researcher at Heinrich Heine University Düsseldorf since 2023. His research focuses on (populist) radical right communication, party competition, and the political system of Germany. He has been part of Wahl-O-Mat Research in Düsseldorf since 2018, researching and supporting VAAs.



Stefan Marschall is a professor of political science at Heinrich Heine University Düsseldorf. His research focuses—among others—on the impact of digital transformation on democratic institutions, political communication, and participation, with a particular interest in voting advice applications such as the Wahl-O-Mat. Ph: Susanne Kurz (HHU).

ARTICLE

Open Access Journal 

Televised Influence: Examining Opinion Formation Through Live Completion of a Voting Advice Application

Nikandros Ioannidis  and Vasiliki Triga 

Department of Communication and Marketing, Cyprus University of Technology, Cyprus

Correspondence: Nikandros Ioannidis (nikandros.ioannidis@cut.ac.cy)

Submitted: 2 September 2025 **Accepted:** 12 November 2025 **Published:** 21 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

The influence of candidate cues on voters’ opinion formulation is a subject of ongoing debate in political science. This study examines the 2023 Cyprus presidential elections, when the leading candidates completed the country’s most popular voting advice application live on national television. This unprecedented event created a natural experiment to test how televised disclosure of candidate positions affects alignment between voters and their preferred candidates. Drawing on voting advice application responses across a wide set of policy questions and employing regression discontinuity in time models, we estimate changes in voter–candidate congruence before and after the broadcast. The results show no uniform effect. For Andreas Mavroyiannis, supporters converged toward his progressive and pro-solution profile, particularly on foreign policy, welfare, and Cyprus settlement questions. For Averof Neophytou and Nikos Christodoulides, congruence did not increase consistently; in some cases, divergence emerged, as explicit cues exposed divisions within their heterogeneous or ambivalent coalitions. The findings provide mixed support for the expectation that cues strengthen congruence only on less entrenched issues. Convergence was most evident on novel questions, but divergence also occurred on identity and Cyprus-related issues. The study highlights that candidate cues can foster alignment but can also generate divergence, depending on candidate characteristics, issue polarisation, and coalition heterogeneity.

Keywords

candidate cues; elite influence; public opinion; voting advice applications; voter–candidate congruence

1. Introduction

Candidate cues help voters form opinions under limited information. Rather than reflecting political knowledge, cue-taking shows how citizens rely on trusted elites to navigate complex choices (Lupia, 1994; Popkin, 1991). This behaviour may simplify decision-making but can also weaken independent judgement and accountability (Samuels & Zucco, 2014).

Political cues are signals from elites that help citizens interpret issues, evaluate policies, or position themselves in political debates (Sniderman et al., 1991). Party labels, endorsements, and candidate statements allow voters to align preferences without detailed knowledge. Party cues are considered stable influences because they embody ideological traditions and historical experiences (Campbell et al., 1960). By contrast, candidate cues highlight individuals' roles, especially in systems where personalities dominate competition (Lenz, 2012).

Cues matter most when issues are complex, new, or of low salience, since voters with limited knowledge or weak predispositions rely on trusted elites. Lupia (1994) showed that endorsements in a Californian referendum helped voters make decisions consistent with complete information. Cohen (2003) found that people often follow their party's stance even when it conflicts with ideology. Druckman et al. (2013) demonstrated that partisan polarisation amplifies cue effects by shaping the interpretation of conflicting messages. Yet, cues are weaker when attitudes are crystallised, salience is high, or credibility is lacking (Bullock, 2011; Nicholson, 2012). Cues may also shape preferences, as voters adjust to their candidate's positions rather than select candidates based on prior congruence (Druckman, 2001; Lenz, 2012).

The present study examines these dynamics in the 2023 presidential elections in Cyprus. For the first time, the three leading candidates (Nikos Christodoulides, Andreas Mavroyiannis, and Averof Neophytou) completed the country's most widely used voting advice application (VAA) live on national television. VAAs are now a familiar part of European campaigns and can influence vote choice and participation (Garzia & Marschall, 2012; Garzia et al., 2017; Gemenis, 2013). In Cyprus, VAAs had already gained wide popularity, with many citizens using them to assess alignment with candidates or parties.

The 2023 event was exceptional because the three main candidates became VAA users themselves in real time on live broadcast television. Each answered the same 22 policy questions posed to the public. Their responses were immediately recorded, reported by the media, and circulated on social platforms. This created a moment of high visibility in which voters could observe their preferred candidate's stance. It also provided unusual standardisation: all three answered the same questions under the same conditions, reducing contextual biases that complicate studies of cue effects in natural campaigns.

The Cypriot case is well-suited to investigating cue dynamics. First, the televised format guaranteed salience. Voters could directly witness elite positions or encounter them through extensive reporting. Second, the timing created a natural cut-point for quasi-experimental analysis. By comparing VAA responses before and after the broadcasts, it is possible to test whether voters shifted their answers to align more closely with their candidate. Third, the breadth of the 22 questions, spanning social, economic, foreign, and Cyprus-related issues, enables analysis of whether cue effects vary across domains. This design takes advantage of these features. We collected all VAA responses during the campaign and merged them with precise timestamps. For each candidate, we identified the broadcast moment and analysed supporter

responses within windows before and after the event. This allows for event-study models estimating changes in voter–candidate distance. Exploiting this temporal cut strengthens claims of causality, even if it cannot replicate laboratory control. Comparing otherwise similar voters before and after the cue, while adjusting for confounders, provides leverage for attributing shifts to candidate signals.

The scope of the study is twofold. Empirically, it asks whether voters adjusted their positions after their candidate’s televised VAA completion. Theoretically, it tests propositions about when cues influence opinion formation. The findings contribute to debates on elite influence and the role of heuristics in shaping electoral behaviour.

2. Literature Review

Existing research has long established that most citizens rely on shortcuts when forming opinions and making choices. Party labels, elite endorsements, and candidates’ stated positions operate as cues that can guide preferences with relatively low cognitive effort. Foundational accounts describe party and leader cues as efficient heuristics that help voters economise on information costs (Campbell et al., 1960; Downs, 1957; Popkin, 1991). Empirical literature subsequently demonstrated that elite cues can alter attitudes, often substantially, when voters are inattentive (Lupia, 1994; Sniderman et al., 1991). Recent evidence shows that this mechanism remains robust in contemporary settings, including multiparty European democracies (Adams et al., 2012; Arceneaux, 2008; Bullock, 2011; Grewenig et al., 2020).

Across many designs, the central finding is consistent: When a trusted party or leader signals a position, co-partisans and supporters tend to move toward that position, whereas signals attributed to disliked out-groups can push voters away, deepening polarisation (Cohen, 2003; Druckman et al., 2013; Nicholson, 2012). Observational and field-based studies similarly find that cues help voters resolve uncertainty at low cost and that the direction of movement generally follows the source’s identity (Arceneaux, 2008; Bullock, 2011). Yet, as Samuels and Zucco (2014) demonstrated, such responsiveness often reflects partisan loyalty rather than informational updating: citizens align with their preferred leaders even when doing so contradicts their prior views.

The literature also highlights important contingencies. Cue effects depend on context and on the nature of the issue. When motivation to process information is low, citizens are more likely to rely on peripheral cues such as source labels rather than scrutinising arguments (Druckman, 2001). Conversely, when the issue is deeply salient or tied to core values, individuals are more likely to engage in effortful processing and to discount cues that conflict with prior beliefs (Lau & Redlawsk, 2001; Sniderman et al., 1991).

As personalisation has intensified, leader images and perceived competence increasingly guide vote choice, with cues serving as shorthand for anticipated policy direction (Balmas et al., 2014; Dalton & Wattenberg, 2000; Garzia, 2014). Recent work has further demonstrated that personalised leader evaluations increasingly structure voter alignments across the continent (Aarts et al., 2013; Costa Lobo & Curtice, 2014).

Issue characteristics also moderate cue effects. Work distinguishing “easy” symbolic issues from “hard” technical ones shows that elites can more readily structure mass opinion on the latter because prior attitudes are weaker and knowledge is thin (Carmines & Stimson, 1980; Lau & Redlawsk, 2001). Elite

influence also depends on issue salience: cues matter most early in an issue's life cycle, before attitudes crystallise (Chong & Druckman, 2007; Druckman et al., 2013; Slothuus & de Vreese, 2010). Studies of motivated reasoning reinforce this asymmetry: When issues implicate identity, citizens selectively accept congenial elite messages and discount dissonant ones; when issues are peripheral to identity, credible cues can move opinions more freely (Druckman et al., 2013; Nicholson, 2012).

Against this background, VAAs have become a prominent element of electoral campaigns. VAAs ask users to respond to a battery of policy statements and then compute the closeness between the user and parties or candidates. Experimental and quasi-experimental research provides evidence that VAA exposure can update preferences and choices: undecided users may revise their party evaluations in the direction of the recommended match (Garzia et al., 2017) or even be mobilised to participate in elections (Gemenis & Rosema, 2014; Germann & Gemenis, 2019). In line with the spatial voting model (Enelow & Hinich, 1984; Tomz & Van Houweling, 2008), VAAs provide personalised proximity information that reduces informational costs, enabling voters to align with ideologically proximate parties. Yet effects are conditional. Some studies find null or modest impacts on proximity voting and switching (Enyedi, 2016; Mahéo, 2017; Munzert et al., 2020), while others suggest that congruence information broadens voters' consideration sets without altering final vote choice (Ioannidis, 2025). Additionally, in candidate-centred presidential races, publicly revealing a candidate's answers to a VAA collapses uncertainty about that candidate's positions. That moment can function as a concentrated cue shock, potentially producing a measurable increase in congruence between voters and candidates across the contested issues, particularly in contexts where voters lack stable predispositions or where issues are less entrenched in political competition (Druckman, 2001; Lenz, 2012; Tromborg & Albertsen, 2023).

Bringing this literature together suggests the conditions under which a televised candidate-completed VAA should produce measurable shifts in voter–candidate congruence. The cue is credible and salient because it comes directly from the preferred candidate in a public setting. The broadcast timing creates a clear before-and-after window that concentrates attention and reduces ambiguity about when the cue was delivered. In a multiparty presidential contest, some issues will be weakly structured among a candidate's supporters; on those issues, the candidate's stance can coordinate the group and compress dispersion (Slothuus & de Vreese, 2010; Sniderman et al., 1991).

This body of evidence provides the theoretical foundation for two hypotheses: that overall congruence among a candidate's voters increased after the televised VAA completion and that congruence increased most strongly on issues where those voters previously held heterogeneous positions.

The first expectation is straightforward; overall voter–candidate congruence will rise after the televised VAA completion. When candidates publicly reveal their positions across the VAA's policy set, supporters learn where the leader stands. Given the general tendency for voters to align with in-group elite signals, many supporters should adjust their responses toward their preferred candidate, shrinking the distance between their answers and the candidate's answers. In a presidential race with intense candidate visibility, a live, widely covered VAA completion is a straightforward vehicle for such cue-driven updating.

The second expectation concerns where, within the issue set, the movement should be largest. The literature indicates that cue elasticity is greatest where prior attitudes are weak, heterogeneous, or not identity-defining

(Druckman, 2001; Slothuus & de Vreese, 2010). Within a candidate's supporter base, some questions will display high dispersion, signalling uncertainty. For those issues, the candidate's public stance provides a focal point that should concentrate opinions and reduce within-group variance. In other words, congruence should increase most on the issues where supporters were initially most divided.

Finally, unlike typical survey experiments that test hypothetical cues, this study uses a quasi-experimental design embedded in a real campaign, where candidates revealed their positions live on air. Thus, this setting offers higher external validity.

3. Data and Methods

3.1. Choose4Cyprus 2023

The 2023 presidential elections in Cyprus provided a unique political context for studying how voters respond to cues from candidates. These elections were highly competitive and took place in an environment of political fragmentation and disillusionment with traditional politics (Ellinas & Katsourides, 2021; Triga et al., 2024). With incumbent President Nicos Anastasiades' term-limited after two five-year terms, the contest to replace him quickly became a three-way race among the leading candidates: Nikos Christodoulides, Andreas Mavroyiannis, and Averof Neophytou ("Former foreign minister," 2023). Christodoulides, a former foreign minister in the Anastasiades government and long-time member of the Democratic Rally (DISY), ran as an independent but was supported by several centrist parties. Mavroyiannis, a career diplomat endorsed by the Progressive Party of Working People (AKEL), presented himself as a progressive alternative. Neophytou, the leader of the governing DISY party, campaigned on continuity and experience. The election was eventually decided in a runoff on 12 February 2023, when Christodoulides narrowly defeated Mavroyiannis by 51.9% to 48.1%, a margin of under 4%, making it one of the closest contests in recent Cypriot history ("Cyprus politician Christodoulides," 2023; Triga et al., 2024).

Amid this tense campaign, an unusual event took place that shaped public debate and provided a near-experimental setting to study political cue-taking. The digital platform Choose4Cyprus, a VAA developed by academics, was introduced just nine days before the first round of the election. It comprised 22 questions covering issues across the economy, society, governance, the Cyprus problem, and foreign policy. Users would indicate their positions, and the system would calculate which candidate's policy stances aligned most closely with their responses. Before seeing their individual congruence results, users were asked to indicate which candidate they intended to vote for. This self-reported voting intention preceded any exposure to their calculated alignment scores.

Despite its late launch, Choose4Cyprus quickly achieved a wide reach. By election day, 10% of the entire electorate had used the application. This was partly driven by the extraordinary media exposure it received on national television. In an unprecedented move, the three leading candidates completed the Choose4Cyprus VAA live on Cyprus's most popular noon information programme, *Alpha Ενημέρωση*, which has a large audience and shapes midday political discourse.

The televised appearances occurred on consecutive days, the week before the election. On 30 January 2023, Andreas Mavroyiannis participated in the live broadcast. Averof Neophytou followed on 1 February 2023, and

Nikos Christodoulides appeared on 2 February 2023 (all participated around 12:40 pm), just three days before the first round of voting. Each session was organised in collaboration with a member of the Choose4Cyprus research team, who joined the show's host to guide the candidates through the process. The format was simple but impactful: candidates responded to each of the 22 policy statements, live on air, explaining their choices to the audience. Once completed, the system produced the candidate's personal VAA result, revealing which political figure most closely matched their expressed positions. For the analysis, each broadcast's exact date and time were used as candidate-specific cut points when defining the *after* variable.

These events provided a setting where candidates with often strategically ambiguous platforms were forced to state their positions on specific policy issues. Presidential contests in Cyprus often revolve around broad appeals, candidate images, and references to the Cyprus problem rather than precise policy commitments (Triga et al., 2018, 2024). Given that the Cypriot presidency combines the roles of head of state and government, candidates' positions carry significant executive and policymaking implications. For this reason, the televised VAA exercises were framed by commentators as an "ideal scenario" to push candidates to clarify their views and allow voters to see them pinned down on concrete choices.

The outcomes of the televised sessions were widely reported and generated public debate. When Andreas Mavroyiannis completed the VAA, the application returned him as the candidate most closely aligned with his own positions, with a very high level of congruence. This unsurprising result allowed him to present himself as coherent and consistent. Mavroyiannis made light of the outcome, remarking that "for good luck," it was best that the tool advised him to vote for himself. Averof Neophytou experienced a similar outcome when he took the quiz two days later. The VAA recommended him as his own best match, which he embraced, using the opportunity to highlight his commitment to digital innovation ("Choose4Cyprus," 2023; "Poio upopsifio edeikse," 2023).

The most notable and controversial moment came during Nikos Christodoulides' appearance on 2 February 2023. When he completed the VAA, the system did not return him as his own closest match. Instead, the application indicated that Christodoulides' answers aligned more closely with those of Andreas Mavroyiannis than with his own declared platform. This surprising outcome was displayed live on television, just days before the election. Christodoulides appeared somewhat taken aback but responded with humour, noting that the result showed he was close to all three major candidates and that this might be interpreted as evidence of his broad appeal. Nonetheless, the symbolism was striking: the frontrunner in the election was told by an ostensibly neutral tool that he should vote for his main rival ("Christodoulidis: "Eimai aisiodoksos," 2023; "O Christodoulidis dokimase," 2023).

This incident quickly became headline news across the media and social platforms. Newspapers and news sites reported that Christodoulides "matched more with Mavroyiannis than with...himself" ("O Christodoulidis dokimase," 2023), emphasising the moment's irony and raising questions about the clarity of his policy positions. The videos of the televised VAA completions circulated widely on social media, gaining thousands of views on YouTube and Facebook within days. The episode fed into ongoing debates about whether Christodoulides' campaign was too vague, with some commentators arguing that his strategy of remaining ideologically ambiguous allowed many voters to project their preferences onto him. The VAA results seemed to illustrate this ambiguity in a concrete and memorable way (Constantas, 2023; "O Christodoulidis dokimase," 2023).

The televised VAA episodes thus became one of the most discussed features of the final campaign week. This case provides a good opportunity to study voter behaviour in response to elite cues. The unique combination of a widely used VAA, real-time candidate participation on national television, and large-scale subsequent citizen engagement creates conditions resembling a natural experiment. The exact timing of the televised completions is precisely known, which allows for before-and-after comparisons of user responses. The VAA dataset contains thousands of individual-level observations, with timestamps indicating when each user completed the questionnaire. This enabled us to examine whether voters who participated after their preferred candidate appeared on television showed greater alignment with that candidate's positions than those who took it earlier. Moreover, the data allows for disaggregation by issue, making it possible to explore whether alignment increased more on specific questions.

3.2. Methodological Framework

The empirical design of this study exploits the unique timing of the VAA completions, which serve as discrete interventions in the information environment. These moments are treated as discrete interventions in the information environment, creating natural experiments for studying how candidate cues affect voter–candidate congruence. To identify these effects, the study adopts a design conceptually similar to a regression discontinuity in time (RDiT) framework, operationalised as a binned event-study. In RDiT, time itself functions as the running variable, with the cut point being the broadcast moment, and the assumption is that all other factors evolve smoothly across that threshold. Any discontinuous change in outcomes at the cutoff can therefore be attributed to the intervention (Hausman & Rapson, 2018). This design is particularly appropriate in this case because the treatment is time-bound and applies, at least potentially, to the entire electorate simultaneously.

Unlike a sharp discontinuity, however, not all voters are exposed to the treatment simultaneously. Some may have been watching the broadcast, while others only encountered the information later through news coverage or social media. This corresponds more closely to a fuzzy discontinuity, in which crossing the temporal threshold sharply increases the probability of treatment without perfectly determining it (Imbens & Lemieux, 2008). Despite this imperfection, the broadcast marks a moment when the information environment changed for all voters, raising the likelihood of exposure in the hours and days immediately after.

In this study, the event-study specification is used to visualise short-term dynamics before and after each broadcast, showing whether voter–candidate congruence trends remain stable prior to treatment. The RDiT framework, by contrast, provides the formal identification strategy by estimating the discontinuous change in congruence at the precise broadcast cut-off using local regressions on either side. The two approaches are thus complementary: the event study offers a descriptive diagnosis of temporal adjustment, while the RDiT isolates the instantaneous effect at the threshold. This combined design is preferable to a sharp regression discontinuity design because exposure to televised cues is inherently probabilistic rather than deterministic: voters differ in when and how they receive the information. Treating the broadcast as a fuzzy, time-based discontinuity, therefore, better captures the real diffusion of candidate cues through media and public discussion, while still preserving the quasi-experimental logic of a discontinuous change in the informational environment. Nevertheless, we also estimated sharp regression discontinuity models using local polynomial regressions for each candidate. These models employed a triangular kernel, a msd bandwidth selector,

and a first-order local linear specification. The corresponding results, presented in the Supplementary File Table A4, are statistically insignificant (as expected) since most voters did not receive the televised information exactly at the temporal threshold but rather in the following hours and days through indirect exposure.

The analysis employs a binned event-study specification centred on the broadcast cut points. Time was divided into six-hour bins covering the 48 hours before and 48 hours after each event. Each bin is represented by a dummy variable, with the final pre-event bin omitted as the reference category. The coefficients, therefore, capture deviations in voter responses in each time window relative to the immediate pre-treatment period. This specification has two advantages. First, it provides a flexible way to trace the temporal dynamics of cue-taking, rather than imposing a single break. Second, it offers an internal placebo test: coefficients in the pre-event bins should be statistically indistinguishable from zero if the identifying assumptions hold, since no treatment should occur before the broadcast. This makes it possible to detect both anticipatory effects and post-event adjustments, while preserving a transparent baseline for interpretation (Schmidheiny & Siegloch, 2023).

The outcome of interest is voter–candidate congruence, operationalised as an affinity score. For each respondent, the absolute difference between their answer and their chosen candidate’s position was calculated for each of the 22 policy questions in the VAA. These item-level distances are then averaged to produce an overall affinity score, where lower values indicate closer alignment and higher values denote greater divergence. Candidate positions are taken from their televised VAA completions, ensuring that the congruence measure reflects the cues that were disclosed to the electorate during the broadcast. This is crucial to the logic of the design, as the televised responses constitute the informational treatment under examination.

The regressions also include covariates for demographics and political interest to improve precision and guard against compositional changes in the sample. These control variables include age, gender, education, and self-reported political interest. Including them in the analysis helps ensure that the estimated discontinuities truly reflect cue-taking effects, rather than shifts in the composition of respondents (Delli Carpini & Keeter, 1996; Lau & Redlawsk, 2001).

The regression model can be expressed as the outcome (affinity score) regressed on the set of bin indicators and the covariates. This is equivalent to a saturated interrupted time-series design, where each coefficient represents the deviation of congruence in that bin relative to the baseline period just before treatment. When plotted, these coefficients produce the familiar event-study graph. If cue-taking occurs, one would expect to see a stable pre-event pattern, followed by a sharp downward shift in affinity scores (i.e., improved congruence) beginning in the bin immediately after the broadcast and persisting for subsequent bins. The six-hour resolution is chosen deliberately, allowing the analysis to capture immediate and delayed adjustments. Since most voters would only have been exposed to the cues after work, via news articles or online clips, it is plausible that the most substantial changes appear 6–12 hours or even a day later, rather than instantly. The binning framework allows this lagged effect to be visualised and tested directly.

In addition to the main event-study design, the study undertakes supplementary analyses to better understand the structure of voter–candidate congruence and to identify which issues are most susceptible to cue-taking.

This part of the analysis focuses on the issue level, examining how the degree of polarisation within each policy item shapes the scope for change. A weighted bimodality coefficient is calculated for all 22 items prior to any televised broadcasts, capturing the extent to which responses form two opposing camps while correcting for one-sided skew. Issues with higher polarisation scores correspond to sharper divisions within the electorate, while lower scores denote broad consensus. The expectation is that cue-taking effects are strongest on highly polarising issues, where uncertainty and disagreement create greater scope for televised exposure to influence perceptions of candidate-voter alignment.

4. Results

We begin the presentation of the results with a descriptive exploration of the data to establish a baseline understanding of how voter–candidate congruence varies before and after each of the three presidential candidates publicly completed the VAA. Table 1 presents the descriptive statistics. For each of the three candidates, the table reports the mean and variance of affinity scores before and after their respective cut points and the mean categorical age, gender composition, mean education, mean political interest, and number of observations.

Table 1. Summary of affinity scores and demographics by candidate and period.

Candidate	Period	Mean affinity	Variance affinity	Mean age (1–7)	Female (%)	Male (%)	Mean education	Mean political interest	N
Christodoulides	Before	1.41	0.08	3.03	44.09	55.91	5.09	2.94	2,251
Christodoulides	After	1.43	0.08	3.00	50.49	49.51	5.03	2.78	2,384
Mavroyiannis	Before	0.94	0.12	2.81	33.43	66.57	5.20	3.32	1,221
Mavroyiannis	After	1.08	0.14	2.87	45.71	54.29	4.98	3.04	4,322
Neophytou	Before	1.31	0.10	2.87	32.53	67.47	5.15	3.16	1,896
Neophytou	After	1.37	0.10	2.86	38.87	61.13	5.07	3.02	3,692

The results show a consistent pattern across the three candidates: mean affinity scores are slightly higher after the cut point than before. For Christodoulides, the mean affinity increases from 1.41 to 1.43, while for Mavroyiannis it rises more substantially from 0.94 to 1.08. Neophytou’s supporters also display an increase from 1.31 to 1.37 after the cut point. The variances remain relatively stable, with no sharp changes in dispersion. These descriptive results may suggest that, on average, voters appeared slightly further from their preferred candidate after the VAA was completed on live television.

However, caution is needed when interpreting these preliminary figures. One plausible explanation for the modest post-cut point increase in affinity scores is related to the dynamics of participation. Toward the end of the period, when the VAA remained open and received high public attention, a surge of late participants entered the sample. These users were likely less politically interested or less informed, and consequently, their answers may have been less consistent with their preferred candidate’s positions, raising the average affinity score. In this sense, the increases in affinity observed descriptively might not reflect genuine persuasion effects from the televised events but instead changes in the composition of the user base. It is precisely because of this limitation that we move beyond descriptive statistics and employ an RDIT design. The event-study framework

allows us to isolate changes in affinity at the precise cut points of candidate participation, while accounting for temporal dynamics.

It is also notable that the VAA user base differs from the general population in predictable ways. Users were, on average, younger, more educated, and more politically engaged than the Cypriot electorate as a whole. The mean age across all candidate groups falls within the 35–44 category, while the average education level is situated at the higher end of the scale. This pattern is consistent with prior research on VAA users, which has repeatedly shown that VAA users tend to be younger, more politically interested, and better educated than the general electorate (Marschall & Schmidt, 2008). These demographic biases do not invalidate the analysis, but they underline that the results should be interpreted as evidence of cue-taking among a politically attentive and relatively well-educated segment of the public, rather than as a direct reflection of the broader Cypriot electorate. Of the 15,771 users who selected one of the three main candidates, 14,078 (89.3%) provided complete socio-demographic information. Prior to analysis, the dataset was cleaned to remove duplicate entries (multiple completions from the same IP address), speeders, and abnormal response patterns (giving more than 7 consecutive times the same response). These exclusions were applied before arriving at the final sample of 15,771 users who selected one of the three main candidates.

We now turn to the main empirical results. The analysis focuses on changes in the affinity score between voters and their preferred candidate in the period surrounding the televised completion of the VAA. Affinity is measured as the average absolute distance between a user's policy responses and those of their chosen candidate, such that higher values indicate greater incongruence. The design exploits six-hour bins around each broadcast, with the $[-6,0)$ interval as the reference category. All models include controls for age, sex, education, political interest, and hour-of-day and day-of-week fixed effects. Standard errors are clustered by day to account for temporal autocorrelation in survey activity.

Figure 1 plots the estimated coefficients and 95% confidence intervals from the binned models. Each panel corresponds to one of the three leading candidates. The coefficients represent the change in mean affinity relative to the pre-broadcast bin. For space reasons, the full model estimates are reported in Table A1 in the Supplementary File. The results indicate heterogeneity across candidates. As a robustness check, we also replicate the analysis using a 24-hour window with four four-hour bins; results are reported in Table A2 in the Supplementary File. The results with the alternative specifications are similar.

For Mavroyiannis, coefficients prior to the broadcast are strongly positive, suggesting that his supporters were on average further away from his stated positions in the days leading up to the broadcast. Immediately after the broadcast, however, the coefficients turn sharply negative, with significant drops between 12 and 36 hours after the event. This pattern suggests that exposure to his positions helped to reduce the distance between him and his supporters, producing a short-term convergence effect. One possible explanation is that Mavroyiannis was less well-known outside diplomatic circles. The broadcast gave him a rare high-visibility opportunity to state clear positions on salient issues, which may have reassured and aligned his centre-left support base, especially voters on the right of AKEL who were still consolidating behind him. Additionally, the pre-event upward trend for Mavroyiannis appears to reflect composition-related variation in early user activity. When the same event-study specification is replicated under a strict nearest-neighbour matching procedure (see Figure A1 in the Supplementary File), this pre-event pattern is substantially diminished, confirming that compositional differences have no meaningful impact on the estimated effects.

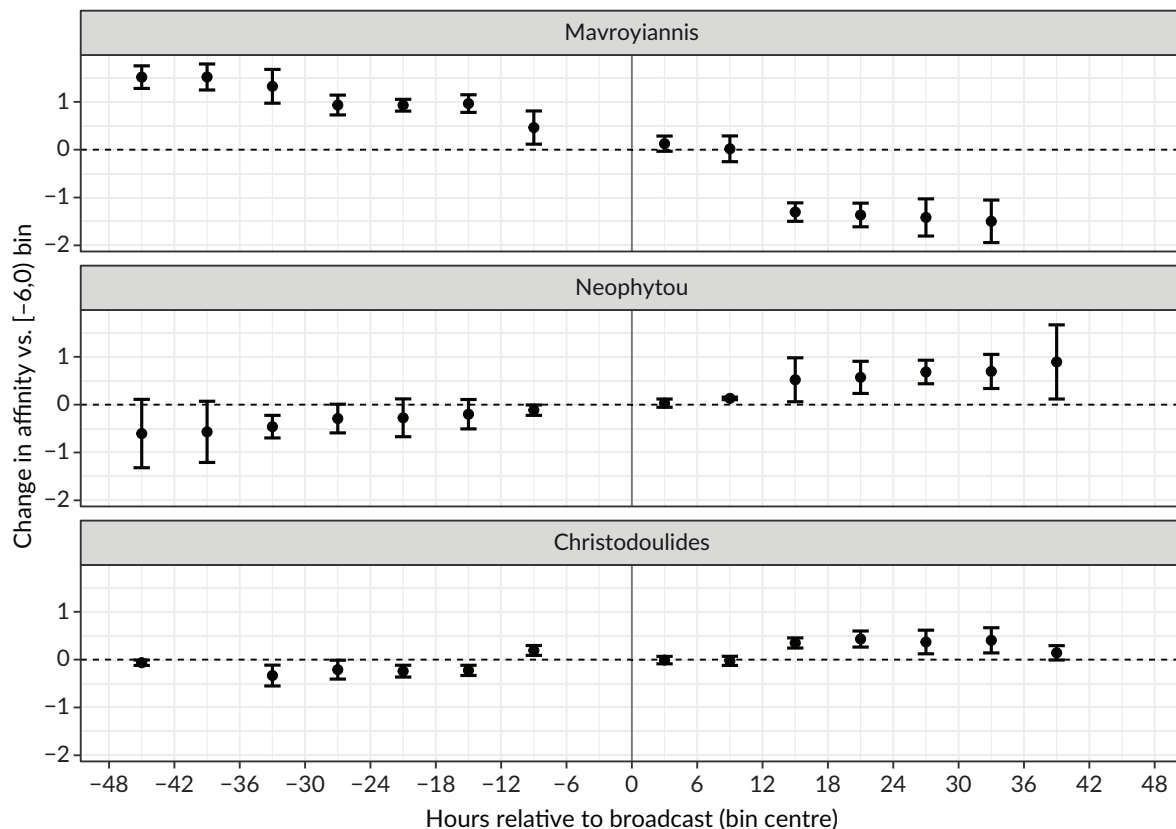


Figure 1. Event-study estimates of change in affinity scores by candidate (bin size: six hours, reference = $[-6,0)$ bin). Notes: Models include controls for age, sex, education, and political interest, plus time dummies; standard errors clustered by day.

Moreover, the matched model reinforces the presence of a clear downward shift beginning in the evening of the broadcast day and continuing into the following day, consistent with genuine post-event convergence rather than artefactual pre-trend noise.

In contrast, Neophytou's results reveal an opposite dynamic. His coefficients are mostly negative prior to the broadcast. After the broadcast, the coefficients turn positive and significant, especially 12 to 36 hours after, indicating that his voters' answers moved further away from his stated positions. Clarifying his positions increased perceived distance on specific issues, eroding rather than strengthening congruence. A possible reason is that during the broadcast, he adopted more centrist positions on welfare and the economy and expressed more liberal views on rights, which may have alienated parts of his traditionally conservative base. In contextual terms, Neophytou faced an uphill battle, as his campaign was overshadowed by intra-party divisions. Many traditional DISY voters were ambivalent, with some leaning towards Christodoulides, who had previously served in DISY governments. The broadcast may have highlighted these internal tensions by making Neophytou's stances appear less representative of the broader DISY electorate, thus widening the gap between him and his declared supporters.

The pattern for Christodoulides is more muted. His coefficients hover close to zero before and immediately after the broadcast, suggesting stability. Yet starting around 12 hours post-broadcast, the coefficients rise significantly, implying that his supporters also drifted somewhat further away from his declared answers.

A plausible explanation is the well-publicised VAA recommendation that matched him with Mavroyiannis rather than himself, which may have prompted some reassessment among his base. Another factor was Christodoulides' strategy of choosing the non-directional answer "neither agree nor disagree" on four key valence issues—environmental protection, same-sex adoption, refugee benefits, and wealth redistribution. Neither Mavroyiannis nor Neophytou resorted to this option in any of the 22 items, regarding it as a signal of dishonesty, but Christodoulides' use of it reinforced perceptions that he was obscuring his positions. More broadly, he adopted a mainstream and pragmatic tone that allowed him to attract support ranging from the centre-left to the right. This catch-all coalition of voters meant that he was already the least congruent candidate with his electorate (Triga et al., 2024), as also seen in Table 1, and therefore, there was little scope for the televised VAA completion to have an effect.

To address potential compositional differences between users who completed the VAA before and after the televised broadcast, we conducted two additional robustness checks replicating the main event-study specification under alternative weighting schemes. First, we implemented a nearest-neighbour matching procedure (3:1 ratio with replacement) on age, sex, education, and political interest. The same event-study model was then re-estimated on the matched data using the matching weights. Second, we performed a raked event-study, in which observations from the reference (pre-broadcast) period were iteratively weighted so that the marginal distributions of key demographics aligned with those of the treated (post-broadcast) group. When raking was infeasible, a stabilised inverse-probability-weighted specification served as a fallback. Both checks are conceptually equivalent to the baseline specification but explicitly adjust for observed differences in user composition. The resulting estimates, shown in Figure A1 (matched) and Figure A2 (raked/inverse probability weighting) in the Supplementary File, closely replicate the main event-study results. The matched specification reduces the mild pre-event trend observed for Mavroyiannis in Figure 1 and strengthens the evidence of a clear downward shift beginning on the evening of his televised completion and continuing into the following day, confirming that compositional imbalance does not account for the observed convergence pattern. These results show that televised completion of the VAA produced asymmetric effects across candidates. For Mavroyiannis, the event appears to have aligned his supporters more closely with him, while for Neophytou and Christodoulides, it seems to have revealed incongruences. These dynamics underscore the centrality of context and candidate profiles in shaping the direction of cue effects. Importantly, they also highlight that elite signals do not always consolidate support but can, in some cases, expose underlying divergences.

We now turn to a more granular analysis of voter–candidate congruence, moving beyond the aggregate affinity score. Instead of treating congruence as a single summary measure across all policy items, we assess how it varies across individual issues and how this variation is conditioned by issue polarisation. The central expectation is that televised exposure should be most consequential on divisive issues. When an issue is already consensual, voters require little additional information to locate themselves relative to a candidate; when it is polarising, televised debates may clarify where candidates stand and help voters reconcile their own views with those positions.

To measure issue polarisation, we computed for each of the 22 VAA policy items a weighted bimodality coefficient using all responses recorded before the first televised broadcast. The coefficient combines skewness and kurtosis to capture the extent to which responses form two opposing camps rather than a single mode. The weighting corrects for one-sided skew, giving higher scores only when both poles of

disagreement are clearly represented. Low bimodality coefficient values, therefore, mark issues of broad agreement, while high values identify genuinely polarised topics that divide the electorate.

Using this issue-level measure, we examined whether voter–candidate distance shifts before and after a candidate’s televised appearance depend on issue polarisation. The models include hour-of-day and weekday fixed effects, with standard errors clustered by day. The unit of analysis is the voter–candidate distance on a single issue.

Figure A4 in the Supplementary File plots the predicted voter–candidate distance before and after each broadcast across levels of issue polarisation. Distance increases with polarisation, that is, voters are generally further from their preferred candidate on divisive issues. For Mavroyiannis, as shown in Figure 2, the broadcast is associated with a reduction in distance on the least polarising issues, with little change on highly polarised ones. For Neophytou and Christodoulides, pre- and post-broadcast lines are parallel and overlapping, indicating no effect. These patterns may reflect the candidates’ strategic choice to maintain moderate positions on polarising issues to avoid alienating segments of their coalitions. However, the observed reduction in Mavroyiannis’ voter–candidate distance runs contrary to our initial expectation, as the overall decline in distance documented in the earlier event-study results appears to stem primarily from improved alignment on the least polarising issues rather than clarification on divisive ones.

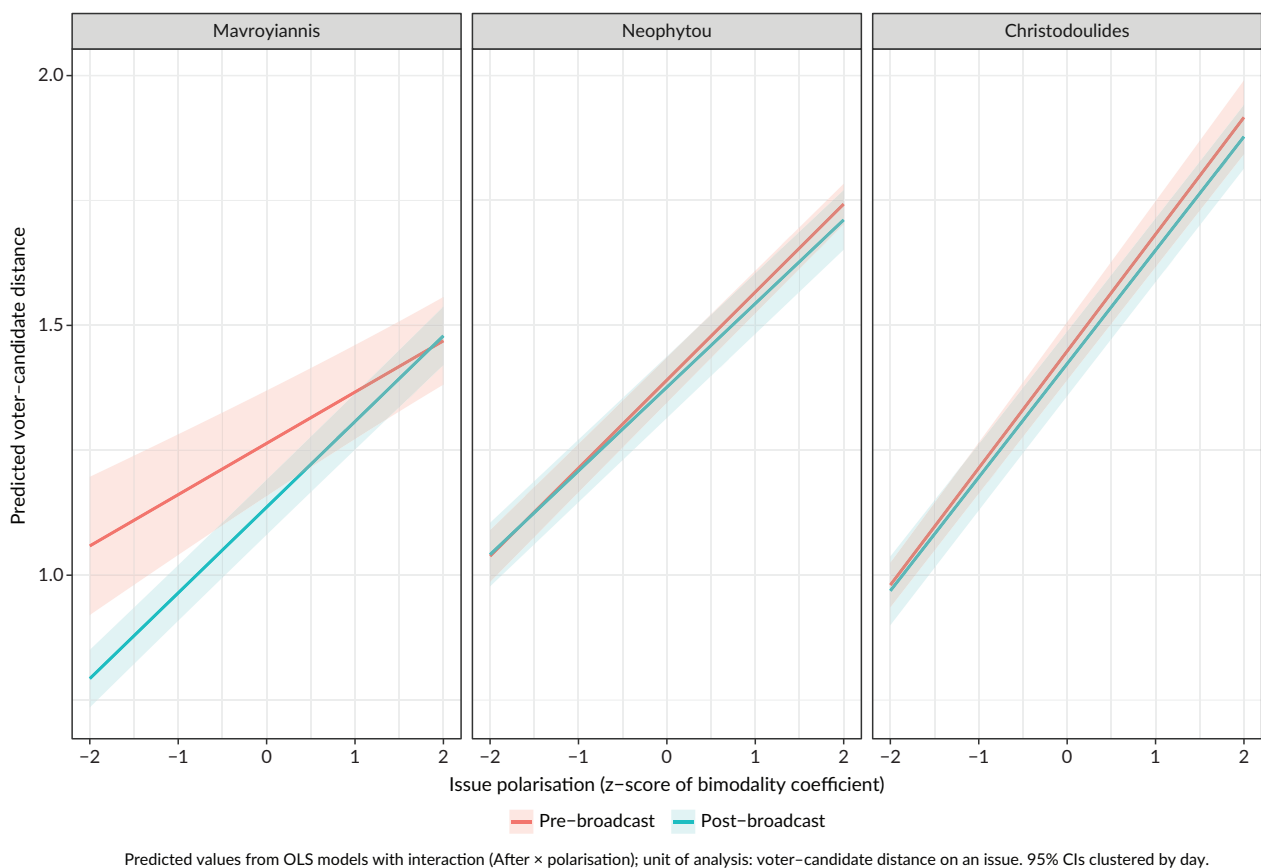


Figure 2. Predicted voter–candidate distance before and after each televised broadcast across levels of issue polarisation.

Finally, the average marginal effects (AMEs) in Table 2 quantify how, after the televised completion, the average distance between a candidate's supporters and that candidate's stance on each item changed relative to the pre-event baseline. Negative values signal convergence, while positive values indicate divergence.

Table 2. Average marginal effects (post-event).

Question	Christodoulides	Mavroyiannis	Neophytou
Environmental protection is more important than economic growth	−0.588***	−0.147	1.194***
The church should have a say in the Minister of Education	0.774***	−0.559	−0.326
Immigrants contribute positively to Cyprus culture	0.482	−1.121***	1.162*
Fence along the Green Line solves immigration	1.724***	−0.307	−2.189**
Same-sex couples equal rights, including adoption	−0.339	−0.946	2.199**
Legalise recreational cannabis	−0.726	0.764	0.410
Abolish refugee benefits if not working	0.287	−1.620***	−0.835*
Accelerate European integration	0.627***	−0.417	0.979***
Support EU sanctions against Russia	−0.795*	−0.693	1.293**
Apply for NATO membership	0.150	−4.322***	−0.412
Privatisation improves service efficiency	0.482	−1.859*	−0.383
Minimal state intervention in the economy	0.608*	0.671	0.164
Redistribute wealth from the rich to the poor	0.183	0.119	−0.169
Increase corporate profit taxes	0.403	0.520*	0.472
Rent-to-buy on state land for housing	0.120	1.757**	−0.777*
“Golden passports” were a sound economic policy	−1.084***	−1.739***	0.496
Greek Cypriots and Turkish Cypriots cannot live together again	1.410*	−3.363**	1.557***
Bi-zonal bi-communal federation is viable	−0.936***	−2.557***	2.164*
Resume negotiations after Crans-Montana	0.328	−2.991***	2.289**
More checkpoints harm the Greek Cypriot side	1.319***	−1.929***	−0.133
Gas pipeline via Turkey before a solution	0.079	0.362	0.320
Referendums are proper decision tools	0.399	−0.672	2.876***

Note: Stars denote statistical significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

For Mavroyiannis, convergence is evident on the core progressive and pro-solution items that defined his campaign. He placed himself firmly at the liberal end on cultural issues (fully supporting same-sex adoption, opposing church influence, fences, and welfare retrenchment). Supporters moved toward him after the broadcast, reflected in negative AMEs such as −1.121 on immigrants' contribution, −1.620 on refugee benefits, and −1.859 on privatisation. On the Cyprus problem, where he unequivocally backed a bizonal federation, resumption of talks, and coexistence, large negative AMEs (−2.557, −2.991, and −1.929) show his base realigning to his pro-federal stance. NATO membership is especially striking: he rejected it outright, and the −4.322 AME indicates that many ambivalent supporters converged strongly toward his opposition. These results fit the expectation that cues matter most where voters are less entrenched and where candidates provide clarity.

For Christodoulides, who adopted centrist positions, the AMEs show selective convergence and divergence. He often placed himself at intermediate points (e.g., “neither agree, nor disagree” on environment vs. growth, “disagree” on church influence, “neither agree, nor disagree” on adoption rights, “disagree” on NATO). This ambiguity broadened his appeal but made him a weak anchor. His supporters converged on some issues associated with the Anastasiades-DISY government, such as the “golden passports” scheme (−1.084) and the bizonal federation (−0.936). Yet divergence appears where his centrism clashed with stronger leanings in his coalition. He opposed the Green Line fence, yet his nationalist supporters held to their harder stance (+1.724). On coexistence with Turkish Cypriots, he expressed support, but rejectionist elements within his base produced divergence (+1.410). In short, Christodoulides’s ambiguity attracted diverse supporters, but the cue exposed their differences: convergence on broadly consensual themes, divergence where middle-of-the-road cues met entrenched predispositions.

For Neophytou, the pattern is bifurcated. He consistently adopted market-liberal, pro-Western, and moderately conservative positions: favouring fences, welfare conditionality, NATO, sanctions on Russia, minimal state intervention, and privatisation. His supporters converged on some of these, notably the fence (−2.189), refugee benefits (−0.835), and privatisation (−0.383), suggesting the cue clarified his stance. Yet substantial divergence also appears, particularly on Europe, social liberalism, and the Cyprus problem. Although he strongly supported a bizonal federation and resumption of talks, his supporters diverged after the event (+2.164 and +2.289). On cultural issues, divergence was pronounced: despite supporting LGBT adoption rights, his base moved sharply away (+2.199), reflecting the disconnect between his progressive stance and the conservative instincts of Cyprus’s mainstream right. These patterns reveal that the televised cue accentuated fissures in his base: alignment improved on traditional centre-right issues of security and markets, but supporters resisted his more cosmopolitan or pro-solution signals.

To further probe the dynamics of cueing, we plot two RDIT models for NATO membership and same-sex couples’ adoption rights. While the AMEs reveal whether congruence increased or decreased overall, these figures allow us to observe when, in the hours following the televised completion, the candidates’ cues began to exert influence on their supporters.

Figure 3 shows how congruence with each candidate’s stated stance on NATO membership evolved in six-hour bins before and after the televised completion. Mavroyiannis, who firmly opposed NATO entry, experienced rapid convergence as his supporters adjusted toward his position in the hours after the broadcast. This shift became primarily apparent 12 hours after the televised completion. Christodoulides, who also opposed NATO membership, saw no systematic movement, suggesting his cue did not alter his divided base on the issue. Neophytou, who clearly backed NATO, also saw no major shifts, consistent with the idea that his pro-Atlantic supporters were already aligned pre-event.

Figure 4 depicts congruence dynamics on same-sex couples’ adoption rights. Mavroyiannis, who strongly supported equal rights, experienced no substantial post-event shifts, as his base was already aligned with his liberal position. Christodoulides, who expressed a more ambivalent stance, likewise, shows little movement, reflecting a coalition split between liberal and conservative instincts. By contrast, Neophytou, who supported same-sex adoption, saw significant divergence: many of his socially conservative supporters moved further away from his declared position after the broadcast, underscoring the difficulty of reconciling his progressive stance with the preferences of his traditionalist electorate. This divergence is again evident after the plus 12-hour bin.

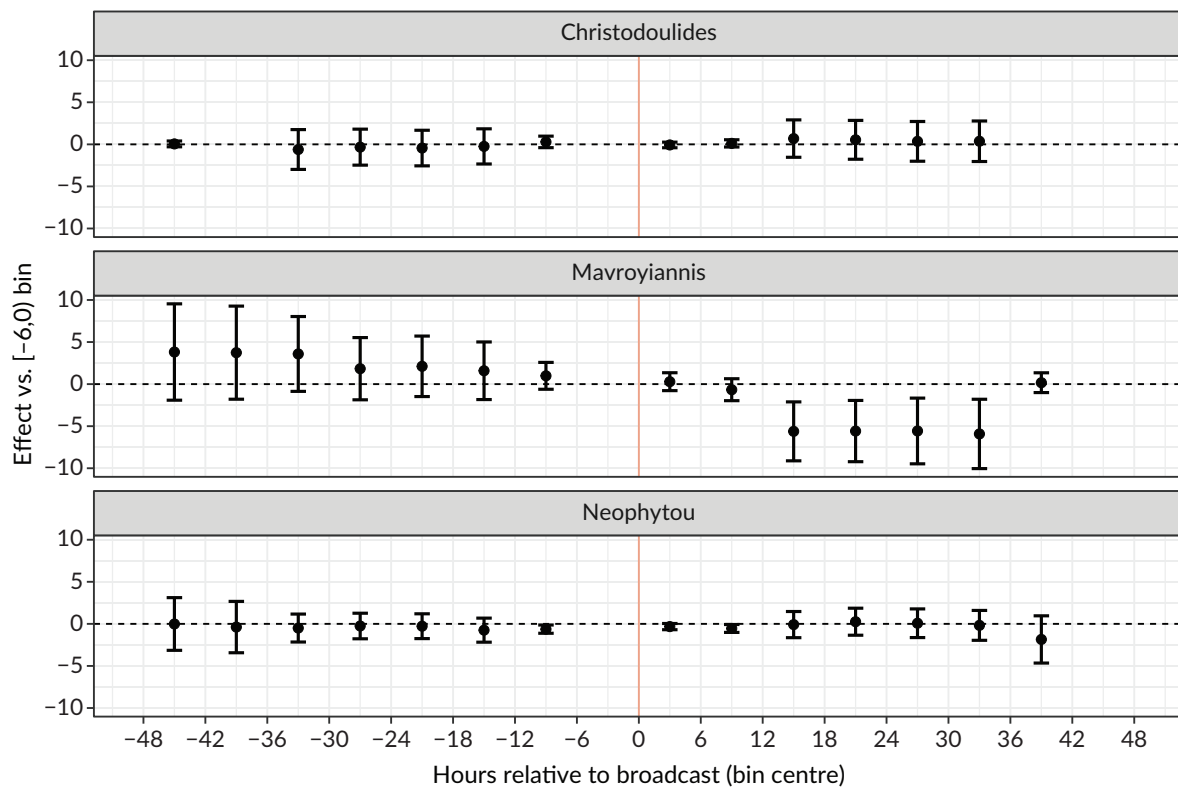


Figure 3. Effect of televised VAA completion on voter–candidate congruence for NATO membership.

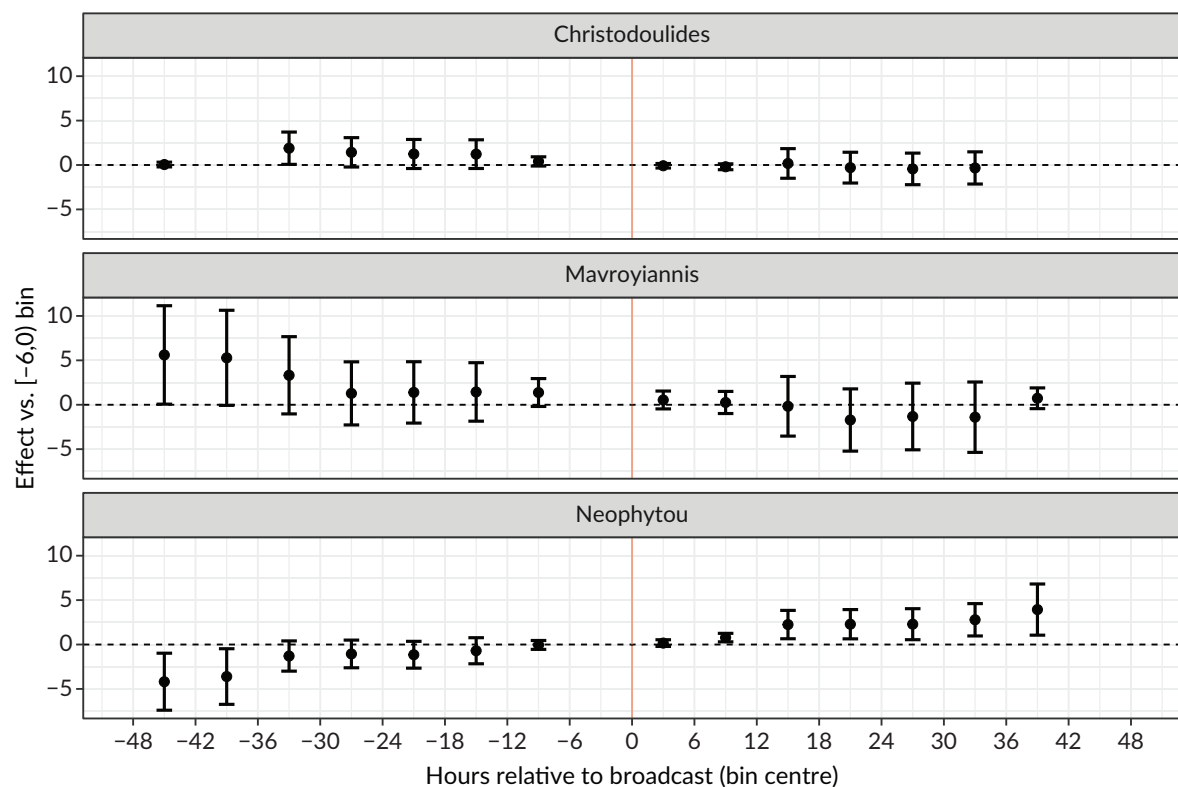


Figure 4. Effect of televised VAA completion on voter–candidate congruence for same-sex couples' adoption.

5. Conclusion

The analysis of the televised completion of the VAA by the three leading candidates in the 2023 Cypriot presidential election provides an unusual window into how candidate cues operate in a high-salience campaign. The first expectation was that congruence between voters and their chosen candidate would increase after the televised event, as supporters aligned themselves more closely with the positions their preferred candidate made explicit. The evidence from the regression discontinuity models does not confirm this expectation in any general sense. There was no uniform pattern across candidates, and the effect of the televised cue was contingent on candidate profiles. For Mavroyiannis, who campaigned with a clear progressive and pro-solution profile, the televised disclosure of positions drew his supporters closer to him, though this convergence was mostly concentrated on low-polarisation issues rather than divisive ones. For Neophytou and Christodoulides, the effect was the opposite: rather than reducing the gap between the candidate and the base, the cue revealed and in some cases, widened underlying divergences. The first expectation is therefore only conditionally supported. Cues can strengthen congruence, but they can also trigger divergence when they expose the heterogeneity of electoral coalitions or highlight the ambiguous positioning of candidates (Bullock, 2011).

This outcome speaks directly to the role of candidate characteristics. Mavroyiannis represented a more ideologically coherent base, largely composed of AKEL voters who expected a pro-solution, progressive stance. His VAA completion confirmed these expectations and clarified his positions on issues such as NATO membership. Supporters responded by adjusting their responses, producing observable convergence. Neophytou, by contrast, led a party with an established pro-Western and market-liberal identity, but his electorate included socially conservative voters and members of a party fractured by the candidacy of Christodoulides. His explicit stance favoring socially liberal positions, such as adoption rights, produced divergence among his supporters. Christodoulides relied on ambiguity and attempted to cultivate a broad appeal across the spectrum. The VAA completion forced him to reveal intermediate or non-committal positions, which failed to anchor his coalition. On issues tied to national identity and the Cyprus problem, his cues even generated divergence. Ambiguity may help in attracting broad support, but the disclosure of concrete positions risks alienating segments of such a coalition.

The second expectation was that congruence would increase mainly on issues less embedded in Cypriot political competition and where voters are more heterogeneous. The supplementary issue-level analysis, using the weighted bimodality coefficient to capture issue polarisation, provides a different view. The results show that congruence is generally higher on consensual, low-polarisation issues, while divisive issues tend to produce larger voter–candidate gaps. Crucially, television broadcasts did not reduce these differences on polarizing issues. Instead, the narrowing of the gap between voters and candidates seems to be due to better alignment on less divisive issues, and this only in the case of Mavroyiannis. This suggests that televised cues primarily clarified already cohesive positions rather than bridging divides. One possible explanation is that all three candidates sought to maintain moderate stances on contentious issues, limiting the scope for convergence among opposing camps.

Taken together, these findings highlight that cue-taking is not a mechanical process of alignment but a conditional one. Candidate characteristics, the internal cohesion of their support bases, and the polarisation of issues all interact to shape outcomes. Rather than a universal effect of televised cues on congruence, the

evidence points to asymmetric dynamics. Candidates with coherent and less established images may use such events to strengthen congruence, while candidates with heterogeneous coalitions or those relying on ambiguity may find that the same events reveal divisions.

The study, therefore, contributes to the literature on VAAs and political heuristics by showing that VAAs do not simply provide information to voters but, when candidates themselves participate, they create highly visible cueing moments with varied consequences. Our findings suggest that when candidates reveal their positions publicly, the effect is mediated by the nature of their coalition and the issues at stake.

The theoretical framework underpinning this study emphasises that cues act as heuristics that help voters resolve uncertainty, but their effectiveness is conditioned by issue polarisation and predispositions. The results are consistent with this framework: cues were effective for Mavroyiannis because his base was predisposed to accept his progressive positions; cues failed for Neophytou and Christodoulides because predispositions diverged from their stances, or the candidates lacked credibility as anchors due to ambiguity. Notably, the study shows that divergence is also a possible outcome of cueing. This challenges the assumption in much of the literature that elite cues primarily increase alignment. When candidates cannot sustain ambiguity, cues can expose and deepen divisions within their coalitions.

The implications extend beyond Cyprus. In multiparty systems with fluid alignments, elite signals may produce asymmetric or contradictory effects. VAAs and similar tools that publicise elite positions can clarify competition but may destabilise coalitions built on ambiguity. For scholars of opinion formation, the lesson is that cues should be studied not as uniform mechanisms but as contingent interventions whose effects depend on both candidates and voters.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The data underlying this article are available from the authors upon reasonable request.

LLMs Disclosure

Large Language Models were used solely for proofreading and minor editorial improvements to the final manuscript. Specifically, ChatGPT 5.0 was employed for this purpose.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Aarts, K., Blais, A., & Schmitt, H. (2013). *Political leaders and democratic elections*. Oxford University Press.
- Adams, J., Ezrow, L., & Leiter, D. (2012). Partisan sorting and niche parties in Europe. *West European Politics*, 35(6), 1272–1294.
- Arceneaux, K. (2008). Can partisan cues diminish democratic accountability? *Political Behavior*, 30(2), 139–160.
- Balmas, M., Rahat, G., Sheafer, T., & Shenhav, S. R. (2014). Two routes to personalised politics: Centralised and decentralised personalisation. *Party Politics*, 20(1), 37–51.

- Bullock, J. G. (2011). Elite influence on public opinion in an informed electorate. *American Political Science Review*, 105(3), 496–515.
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. University of Chicago Press.
- Carmines, E. G., & Stimson, J. A. (1980). The two faces of issue voting. *American Political Science Review*, 74(1), 78–91.
- Chong, D., & Druckman, J. N. (2007). A theory of framing and opinion formation in competitive elite environments. *Journal of Communication*, 57(1), 99–118.
- Choose4Cyprus: “Psifise Averof” eipe ston Averof to chrisimo ergaleio (Video). (2023, February 1). *AlphaNews*. <https://www.alphanews.live/politics/choose4cyprus-psifise-averof-eipe-ston-averof-to-chrisimo-ergaleio-vinteo>
- Christodoulidis: “Eimai aisiodoksos gia to apotelesma ton eklogon” (Video). (2023, February 2). *AlphaNews*. <https://www.alphanews.live/politics/christodoulidis-eimai-aisiodoxos-gia-to-apotelesma-ton-eklogon-vinteo>
- Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85(5), 808–822.
- Constantas, K. (2023, February 8). To ergaleio pou taraxe ta nera tis proeklogikis. *Limassol Today*. <https://limassoltoday.com.cy/l-air-uncut/to-ergaleio-pou-taraxe-ta-nera-tis-proe>
- Costa Lobo, M., & Curtice, J. (2014). *Personality politics? The role of leader evaluations in democratic elections*. Oxford University Press.
- Cyprus politician Christodoulides wins presidential vote. (2023, February 12). *Reuters*. <https://www.reuters.com/world/middle-east/cypriot-diplomats-face-off-cliffhanger-presidential-vote-2023-02-12>
- Dalton, R. J., & Wattenberg, M. P. (2000). *Parties without partisans: Political change in advanced industrial democracies*. Oxford University Press.
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Downs, A. (1957). *An economic theory of democracy*. Harper.
- Druckman, J. N. (2001). On the limits of framing effects: Who can frame? *The Journal of Politics*, 63(4), 1041–1066.
- Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, 107(1), 57–79.
- Ellinas, A. A., & Katsourides, Y. (2021). The silent electoral earthquake in Cyprus: A crisis of political representation. *South European Society and Politics*, 26(3), 413–436.
- Enelow, J. M., & Hinich, M. J. (1984). *The spatial theory of voting: An introduction*. Cambridge University Press.
- Enyedi, Z. (2016). The influence of voting advice applications on preferences, loyalties and turnout: An experimental study. *Political Studies*, 64(4), 1000–1015.
- Former foreign minister holds lead in Cyprus vote, runoff on Feb 12. (2023, February 5). *Reuters*. <https://www.reuters.com/world/middle-east/former-foreign-minister-holds-lead-cyprus-vote-runoff-feb-12-2023-02-05>
- Garzia, D. (2014). *Personalization of politics in Western democracies: Causes and consequences of leader effects on voting behaviour*. ECPR Press.
- Garzia, D., & Marschall, S. (2012). Voting advice applications under review: The state of research. *International Journal of Electronic Governance*, 5(3/4), 203–222.
- Garzia, D., Trechsel, A. H., & De Angelis, A. (2017). Voting advice applications and electoral participation:

- Evidence from a quasi-experiment during the 2011 Swiss federal elections. *Political Communication*, 34(3), 434–443.
- Gemenis, K. (2013). Estimating parties' policy positions through voting advice applications: Some methodological considerations. *Acta Politica*, 48(3), 268–295.
- Gemenis, K., & Rosema, M. (2014). Voting advice applications and electoral turnout. *Electoral Studies*, 36, 281–289.
- Germann, M., & Gemenis, K. (2019). Getting out the vote with voting advice applications. *Political Communication*, 36(1), 149–170.
- Grewenig, E., Lehwald, S., & Roth, C. (2020). Political polarization and attitudes towards climate change in Europe. *Journal of Public Economics*, 187, Article 104229.
- Hausman, C., & Rapson, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, 10(1), 533–552.
- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615–635.
- Ioannidis, N. (2025). The power of alignment: How personalized information shapes voter decisions. *Journal of Information Technology & Politics*. Advance online publication. <https://doi.org/10.1080/19331681.2025.2512856>
- Lau, R. R., & Redlawsk, D. P. (2001). Advantages and disadvantages of cognitive heuristics in political decision making. *American Journal of Political Science*, 45(4), 951–971.
- Lenz, G. S. (2012). *Follow the leader? How voters respond to politicians' policies and performance*. University of Chicago Press.
- Lupia, A. (1994). Shortcuts versus encyclopedias: Information and voting behavior in California insurance reform elections. *American Political Science Review*, 88(1), 63–76.
- Mahéo, V. A. (2017). Voting advice applications and electoral preferences: The role of political sophistication. *Electoral Studies*, 45, 75–84.
- Marschall, S., & Schmidt, C. K. (2008). Preaching to the converted or making a difference? Mobilizing effects of an internet application at the German general election 2005. In D. Farrell & R. Schmitt-Beck (Eds.), *Non-party actors in electoral politics: The role of interest groups and independent citizens in contemporary elections* (Vol. 8, pp. 259–278). Nomos.
- Munzert, S., Rubba, C., Ramirez, M. D., & Trechsel, A. H. (2020). What drives the effect of voting advice applications? Evidence from a survey experiment. *Political Science Research and Methods*, 8(4), 845–862.
- Nicholson, S. P. (2012). Polarizing cues. *American Journal of Political Science*, 56(1), 52–66.
- O Christodoulidis dokimase to “Choose4Cyprus” kai tairiazei perissotero me Mavrogianni para me ton eauto tou (Video). (2023, February 2). *Politis*. <https://www.politis.com.cy/politis-news/politiki/589118/o-xristodoulidis-dokimase-to-choose4cyprus-kai-tairiazei-perissotero-me-mavrogianni-para-me-ton-eafto-tou-vinteo>
- Poio upopsifio edeikse ston Mavrogianni to erotimatologio Choose4Cyprus (Video). (2023, January 30). *AlphaNews*. <https://www.alphanews.live/politics/poio-ypopsifio-edeixe-ston-mavrogianni-to-erotimatologio-choose4cyprus-vinteo>
- Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. University of Chicago Press.
- Samuels, D., & Zucco, C., Jr. (2014). The power of partisanship in Brazil: Evidence from survey experiments. *American Journal of Political Science*, 58(1), 212–225.
- Schmidheiny, K., & Siegloch, S. (2023). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*, 38(5), 695–713.

- Slothuus, R., & de Vreese, C. H. (2010). Political parties, motivated reasoning, and issue framing effects. *The Journal of Politics*, 72(3), 630–645.
- Sniderman, P. M., Brody, R. A., & Tetlock, P. E. (1991). *Reasoning and choice: Explorations in political psychology*. Cambridge University Press.
- Tomz, M., & Van Houweling, R. P. (2008). Candidate positioning and voter choice. *American Political Science Review*, 102(3), 303–318.
- Triga, V., Ioannidis, N., & Djouvas, C. (2024). The waning of ideology? Presidential elections in the Republic of Cyprus, 5 February 2023. *South European Society and Politics*, 28(2), 177–206. <https://doi.org/10.1080/13608746.2024.2304445>
- Triga, V., Mendez, F., & Djouvas, C. (2018). Post-crisis political normalisation? The 2018 presidential elections in the Republic of Cyprus. *South European Society and Politics*, 24(1), 103–127.
- Tromborg, M., & Albertsen, M. (2023). Elite cues and the conditionality of persuasion. *Political Behavior*, 45(2), 361–385.

About the Authors



Nikandros Ioannidis is a special teaching staff member at the Cyprus University of Technology. His research focuses on political communication, political behaviour, and digital tools in democratic participation, with an emphasis on experimental and computational methods.



Vasiliki Triga is an associate professor at the Cyprus University of Technology. Her research examines political behaviour, digital democracy, and public opinion, with a particular interest in the effects of online tools and technologies on democratic processes and citizen engagement.

ARTICLE

Open Access Journal 

Voting Advice Application Use Increases Party Position Knowledge: An Experimental Study Among Belgian Youngsters

Joke Matthieu ^{1,2} , Laura Jacobs ² , Matthias Van Campenhout ² ,
and Stefaan Walgrave ² 

¹ Department of Political Science, Vrije Universiteit Brussel, Belgium

² Department of Political Science, University of Antwerp, Belgium

Correspondence: Joke Matthieu (joke.matthieu@vub.be / joke.matthieu@uantwerpen.be)

Submitted: 29 August 2025 **Accepted:** 11 December 2025 **Published:** 28 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Political knowledge is vital for democratic citizenship, yet many voters struggle to correctly identify party positions. This article investigates whether voting advice applications (VAAs) can increase party position knowledge (PPK) among young voters. We present results from a large-scale experiment in Flanders (Belgium) using two real-world VAAs: De Stemtest, a generic VAA targeting the general electorate, and De Jongerenstemtest, a youth-oriented version specifically designed to appeal to younger citizens. Respondents aged 16–30 ($N = 2,291$) were randomly assigned to a control group or one of the treatment conditions. Multilevel logistic regression models demonstrate that VAA exposure does not uniformly enhance PPK: the generic VAA produced no significant learning effect, whereas the youth-oriented VAA significantly improved respondents' ability to identify party positions correctly. Contrary to expectations, the learning effects did not vary systematically across parties or between mainstream and niche parties. Our findings provide rare causal evidence of PPK effects from real-world VAA exposure, highlighting the importance of tailoring such tools to specific audiences.

Keywords

experimental design; party position knowledge; political knowledge; survey experiment; voting advice applications; young voters

1. Introduction

Political knowledge is a cornerstone of democratic citizenship. It enables voters to navigate the political landscape, hold political elites accountable, and make informed electoral choices (Barabas et al., 2014; Delli Carpini & Keeter, 1996). Party position knowledge (PPK) refers to citizens' awareness of where political parties stand on specific policy issues. This type of knowledge is particularly important in multiparty systems, where programmatic competition often spans multiple ideological dimensions and voters cannot rely solely on simple heuristics. Voters with knowledge of party positions can engage meaningfully in representative democracies and make electoral choices that align with their substantive policy preferences.

While static political knowledge typically focuses on enduring facts about institutions, leaders, or procedural rules, PPK is policy-specific and dynamic, requiring individuals to track parties' evolving stances across a range of issue domains. As Barabas et al. (2014) argue, policy-specific knowledge imposes high cognitive demands and is often unevenly distributed in the electorate, being concentrated among politically engaged publics. Many voters hold only partial or inaccurate perceptions of party positions, especially in multiparty systems where the number of parties and overlapping issues make it hard to create a clear mental map of the political landscape (Schultze, 2014). Individual-level demand factors such as political interest, cognitive ability, and media use are consistently associated with higher levels of PPK. However, PPK also depends on the availability of information and campaign dynamics (Barabas & Jerit, 2009). In this study, we deal with the latter. We do not deal with individual predictors of PPK but instead focus on a specific type of information that voters may encounter during an electoral campaign: voting advice applications (VAAs).

VAAs are online tools that enable users to compare their own stances on various policy issues with those of the parties on offer and inform them about which parties align more closely with their own policy views. In many countries, VAAs have boomed, and using a VAA has become one of the most employed ways in which voters inform themselves about the party's offer. For instance, in the country we are studying here, Belgium, the most popular VAA during the last election in June 2024 was used more than seven million times, on a voting population of just over eight million. In countries such as the Netherlands, Germany, and Switzerland, similar participation rates are observed.

What VAAs essentially do is inform voters about parties' policy stances, both directly and indirectly. Virtually all VAAs end with an output screen that displays a list of parties along with each party's degree of overlap with, or distance from, the user's own preferences. This way, indirectly, through this comparison with their own preferences, users may learn about what stances parties hold. For instance, a user who gets a high aggregate overlap with a particular party may derive this party's policy positions from that generally high congruence (being the same position as that this voter personally holds). However, some VAAs also allow for the direct learning of parties' policy positions, as they, at the end of the session, provide users the opportunity to dig deeper into the results and compare, policy by policy, their own views with those of each party. If users grab that opportunity, they may directly observe and learn, for each issue, what each party stands for. The VAA, whose effect we study here, did offer such a deepening and direct learning option; however, we have no information on whether our subjects actually utilized this option.

Previous work has already examined the PPK effects of VAA use, but experimental work that can tease out a truly causal effect of VAA exposure remains very limited. Our study offers rare experimental evidence leveraging a real-world VAA while entirely controlling exposure. Moreover, previous work looked at the

effect of VAA use on general users. Here, we examine whether learning effects are different when a VAA is designed for and targeted at a specific population group, namely the young. Indeed, young people may be more sensitive to new information, and a VAA that engages them more, both in terms of substance and format, may promote greater learning effects. Finally, previous work on VAA use and PPK did not investigate whether learning varies across parties; our study looks at party differences and tests whether users learn more about mainstream parties than about niche parties.

Concretely, this article experimentally examines the possible increase in PPK among a sample of more than 2,000 young Belgian voters and compares the PPK effect of participating in a generic “normal” VAA (De Stemtest) with that of the use of a special VAA targeted towards the youth (De Jongerenstemtest). We find that VAA use indeed significantly increases PPK among our young subjects, but that only the specific youth VAA was able to generate a PPK effect; among the young voters in our sample, we do not witness an increase in PPK after completing the generic VAA. Learning of party positions seems to be indistinguishable across all parties, and we do not find that young subjects learn more about some parties’ positions compared to others.

2. Party Characteristics, VAA Use, and PPK

Voters rely on ideological heuristics and party cues to make sense of party positions (Conover & Feldman, 1989; Dahlberg & Harteveld, 2016; Zaller, 1992). These cues are more effective when they are clear and unambiguous (Bischof & Senninger, 2018; Lefevere, 2024). Mainstream, centrist parties frequently adopt more unclear positions, either to appeal broadly, as a result of coalition dynamics and the need for compromise or to avoid electoral backlash (Dalton & McAllister, 2015; Lefevere & Verwee, 2025). Ambiguity—especially in the form of vagueness (the extent to which a policy statement of a party allows for varied interpretations) or inconsistency (referring to a party expressing conflicting positions on an issue)—undermines voters’ ability to identify party positions (Lefevere & Verwee, 2025). For example, some parties adopt clear-cut positions on both the traditional socioeconomic left–right dimension and the newer sociocultural GAL–TAN axis. These parties typically maintain a more consistent ideological profile compared to mainstream parties, aligning as either left-wing or right-wing across a wide range of policy issues. Often, these are radical or niche parties with more ideologically extreme positions and clear ownership over a number of issues (Conover & Feldman, 1989; Jennart, 2025; Lefevere, 2024). This clarity also reinforces the potency of ideological heuristics, particularly on the margins of the ideological spectrum (Dahlberg & Harteveld, 2016; Rovny, 2013). Still, the relationship between party type and PPK might vary across issue domains. Niche parties—such as radical right parties—tend to be most explicit on the issues they “own,” but remain more ambiguous on long-standing economic dimensions (Rovny & Polk, 2020). Conversely, mainstream parties may be easier to locate on these traditional issues, but less so on newer or cross-cutting ones. Nevertheless, we expect that on average, niche parties’ positions are easier to identify because their messages are more homogeneous and less strategically blurred. In the context under scrutiny, niche parties display a high degree of ideological congruence across both economic and cultural dimensions, whereas mainstream parties must manage more internally diverse electorates, resulting in more shifts on these issues. Thus, relying on heuristics should be easier for niche parties than for mainstream parties. Consequently, individuals’ knowledge of party positions before exposure to any treatment—i.e., their baseline or pre-treatment PPK—is expected to be lower for centrist mainstream parties than for more ideologically extreme niche parties (Jennart, 2025).

Hypothesis 1: The baseline level of PPK is higher for niche parties compared to mainstream parties.

VAA became an integral part of electoral campaigns in many European democracies. These online tools enable users to compare their own policy preferences with those of the parties. Hence, VAA hold the potential to serve as a learning tool, helping users build more accurate and comprehensive knowledge of party positions. They expose users to party positions on policy issues (Garzia & Marschall, 2016). VAA not just confront users with the views of parties, they also present the information in a highly structured and accessible format, providing side-by-side comparisons of party stances. This should reduce the cognitive effort required to gather party position information from fragmented bits in the news media (Kamoen et al., 2015). Moreover, VAA confront users with policy areas they might not otherwise engage with, broadening their exposure beyond their usual public. This feature is particularly valuable in multiparty systems, where voters often lack clear heuristics for mapping party positions across multiple ideological dimensions.

Empirical research, employing a variety of methodological approaches, did find a generally positive relationship between VAA exposure and PPK, although the magnitude and consistency of these effects vary. Early work relied primarily on observational designs, either inviting individuals who had just completed a VAA to participate in a follow-up survey (Kamoen et al., 2015) or asking respondents about their past VAA use alongside measures of political knowledge (Schultze, 2014). While these studies often report substantial gains in voters' ability to identify what parties stand for (e.g., Schultze, 2014), this is not always the case (e.g., Heinsohn et al., 2019). Moreover, these designs have been criticized for lacking a proper control group of non-users and for their susceptibility to self-selection biases (Pianzola, 2014), as participants in these studies are often skewed in terms of age, education, and political interest (Heinsohn et al., 2019).

In response to these limitations, scholars have increasingly adopted experimental designs, in which respondents are randomly assigned to either complete a VAA or not, mitigating both problems with self-selection and a lack of a control group. Experimental studies, however, tend to find much smaller effects on PPK, highlighting the challenges of isolating causal impacts (Munzert & Ramirez-Ruiz, 2021). The difference between observational and experimental work thus suggests that part of the observed knowledge gains may be attributed to self-selection.

Hypothesis 2: Exposure to a VAA increases PPK among young voters.

While VAA may enhance PPK in general, most existing tools are designed for a broad electorate and rarely account for the specific needs of younger users. Studies of VAA effects typically evaluate a single generic application aimed at informing the general voting public, which means their findings reflect the average impact across diverse age groups (Munzert & Ramirez-Ruiz, 2021; Schultze, 2014). However, there are reasons to believe that these generic tools may not reach their full potential among younger citizens, who differ from older groups in their political knowledge and media consumption habits.

Younger voters are still shaping their political views and knowledge, often having had less prior exposure to party positions (Neundorff & Smets, 2017). This may make them particularly sensitive to interventions that reduce informational barriers and present content in an accessible way. Moreover, younger audiences typically prefer visually engaging, interactive, and contextually relevant formats that align with their digital media habits (Kivunja, 2014). Generic VAA may therefore be less effective in capturing youngsters' attention or sustaining their interest throughout the tool.

A VAA specifically designed for young voters may address these limitations by incorporating simplified language, age-relevant issue framing, and dynamic design features. These adaptations could increase users' attention and motivation—two critical components of learning (Bandura, 1986; Kivunja, 2014). Higher motivation is expected to encourage younger users to more actively engage with party stances, which is key to transforming exposure into knowledge. Furthermore, such a tool could present information in smaller, more digestible units and provide immediate, tailored feedback on how users' responses align with party positions. Research on gamification, for instance, shows that such design choices—including interactive feedback and visually engaging elements—are particularly effective in helping younger audiences comprehend and retain new information (Hamari et al., 2014). Hence, a youth-oriented VAA may have the potential to foster more active learning among the young.

Hypothesis 3: Exposure to a youth VAA increases PPK among young voters more than exposure to a generic VAA.

In addition to general expectations of VAA effects on PPK, we argue that the magnitude of the knowledge gain may vary systematically depending on party characteristics. We expect the effect of VAA use on PPK to be greater for mainstream parties than for niche parties. This expectation stems from the differential baseline visibility and clarity of cues for party positions. Since mainstream parties' positions are less clear to begin with (see Hypothesis 1), there is more to gain in terms of PPK for their positions compared to those of niche parties. This increases the likelihood that voters already know the latter's positions before using a VAA (Conover & Feldman, 1989; Jennart, 2025; Rovny, 2013). VAAs expose users in a structured and accessible way to a wide range of issue positions and expose "issue publics" (Henderson, 2014; Krosnick, 1990) to issues they may care less about. Such a presentation is especially useful when party positions are unclear or inconsistently communicated elsewhere, making VAAs particularly informative regarding mainstream parties whose stances are less distinctive and more susceptible to misperception or lack of awareness. For niche parties, the potential for knowledge gain through VAA exposure may be lower due to a ceiling effect, where users encounter relatively little new information, and thus, their PPK may not increase significantly post-treatment (Schultze, 2014; Van De Pol et al., 2014).

Hypothesis 4: Exposure to a VAA increases PPK among young voters more with respect to mainstream parties' positions than to those of niche parties.

3. Data and Method

3.1. Case Description

Participants in our study are from Flanders, the Northern part of Belgium—a parliamentary democracy with a multiparty system, high electoral volatility, and party fragmentation featuring both mainstream and niche parties (Deschouwer, 2012). The three Flemish mainstream parties—Christian-democrats (CD&V), liberals (Open Vld), and socialists (Vooruit)—were founded in the 19th century. These parties have declined electorally in favor of newer parties founded in response to newly emerging cleavages (Delwit & Van Haute, 2021; Dennison & Kriesi, 2023; Luybaert, 2025). In this article, we consider the Green Party (Groen), the radical right (Vlaams Belang, VB), and the radical left (PvdA) to be niche parties. Since the beginning of the 2000s, a renewed Flemish nationalist party (N-VA) has also been very successful.

On June 9, 2024, regional, federal, and European elections were held. Belgium has a system of compulsory turnout. For the first time, the voting age was lowered for the European elections: young citizens who had reached the age of 16 were required to vote for the European Parliament, while the minimum age for voting in other elections remained 18 years. This means that around 600,000 voters had to vote for the first time. We focus here on this crucial group of 16–30-year-olds (Generation Z)—both new voters and young citizens with limited electoral experience. Research indicates that civic habits develop later for younger generations (Jennings & Stoker, 2004) and that this age range is crucial for the formation of political behavior (Boonen, 2015). Generation Z often settles down later, delaying milestones such as marriage and parenthood, which in turn further postpones civic engagement. Studying the full 16–30 age span helps capture these shifting patterns. The upper limit of 30 also reflects common age caps for youth party wings and organizations in Belgium.

During the 2024 campaign, the Flemish public broadcaster VRT and the newspaper *De Standaard*, in collaboration with the University of Antwerp and Tree Company (Walgrave et al., 2025), launched *De Stemtest*, a generic VAA, and *De Jongerenstemtest*, a VAA aimed at youngsters. While *De Stemtest* has been well-established since its foundation in 2004, *De Jongerenstemtest* marked the first large-scale initiative tailored specifically to young voters in Flanders. Both VAAs were extensively consulted during the campaign: *De Stemtest* was completed 3,931,582 times and *De Jongerenstemtest* was completed 601,977 times, reaching a large segment of the electorate. In fact, *De Stemtest* set a global record for individual-level usage.

3.2. Experimental Design

We leverage a between-subjects experiment. Data was collected from April 4 until April 14, 2024, during the electoral campaign in the run-up to the June 9 elections. Both VAAs used in this study were only made publicly accessible on April 15. We examine the effect of treatment by real-world VAAs that were not yet publicly available. This means we had complete control over exposure and are sure that subjects were not pre-treated. Other experimental studies on VAAs often rely on “real-life” VAAs that have already been launched (e.g., Enyedi, 2016; Garry et al., 2019; Munzert et al., 2021), but we test the possible implications of a real-life VAA without the possibility that respondents had already participated in it, optimizing the chances of fully isolating the causal effect. The fact that respondents in our study were exposed to the actual VAA, which was later launched in the lead-up to the June 9, 2024 elections in Belgium, adds substantial ecological validity to our study.

Participants were drawn from a five-wave youth panel, which was approved by the ethics board of the University of Antwerp. Our study was part of Wave 2. The panel was recruited using convenience sampling, with researchers and student assistants engaging youth organizations and recruiting participants from schools and other youth locations, relying on a stratified sample of large, mid-size, and small towns across Belgium (Jacobs et al., 2024). A total of 3,051 participants were included in the present study. A total of 1,566 respondents were randomly assigned to the control group, who completed a survey that included both control and dependent variables; 1,485 respondents were randomly assigned to either *De Stemtest* or *De Jongerenstemtest*. These treatment group respondents were first invited to complete a short survey that included only control variables. Then they were invited to participate in a second survey in which one of the VAAs was embedded, and where, right after completing the VAA, they were asked to answer questions

about the main dependent variable, PPK. Respondents received monetary compensation for participating in the experiment and the panel in general. Youth panel members earned €5 for each wave they completed, totaling €50 if they participated in all five waves. Additionally, respondents assigned to an experimental condition were offered an extra €5 on top of their initially earned compensation for participating in the panel. While acknowledging potential negative effects, compensating respondents for their participation has proven effective in reaching typically underrepresented populations (Singer & Kulka, 2002) and increasing the overall representativeness of studies (DeCamp & Manierre, 2016). This strategy proved fruitful, given that the panel attrition between Wave 1 and Wave 2 was only 14.6%. Figure 1 offers an overview of the experimental design.

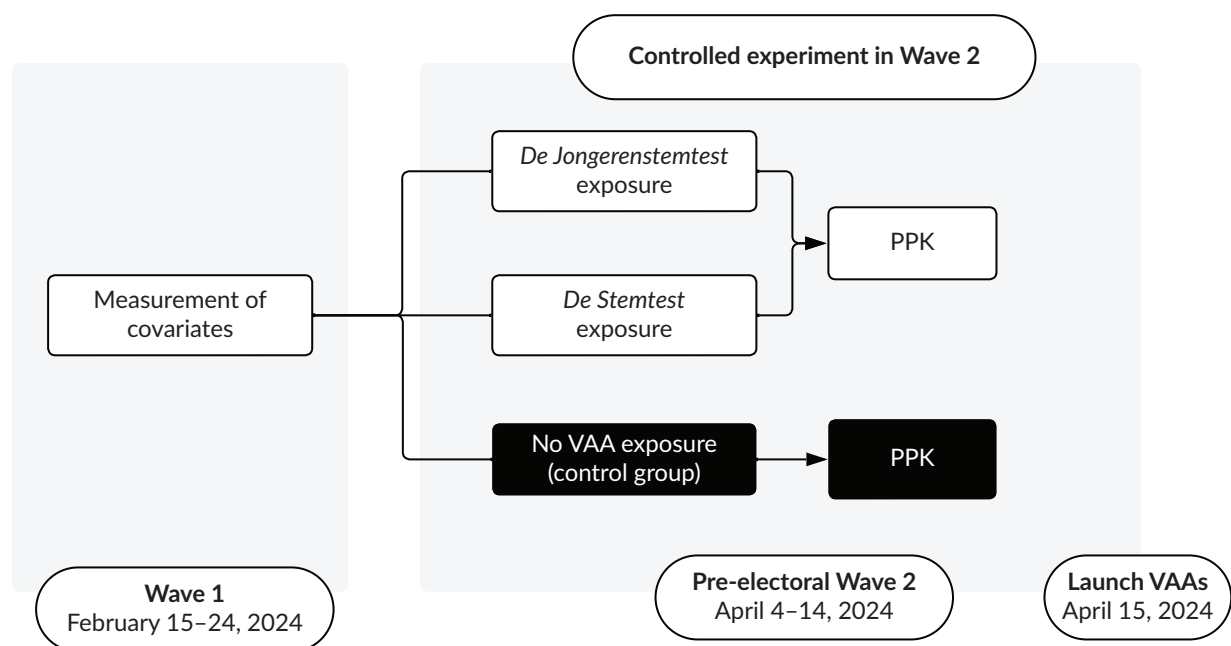


Figure 1. Overview of the experimental study design.

Those in the De Stemtest condition were exposed to a VAA targeting a general audience. Respondents assigned to the De Jongerenstemtest condition participated in a VAA designed specifically for younger citizens, incorporating features that should appeal to youth. The two VAAs differed in terms of statements, the inclusion of status quo information, and overall layout. Each VAA included 35 statements, of which 15 overlapped, and 20 were unique to each version (Walgrave et al., 2025). De Jongerenstemtest featured items more attuned to young citizens' daily life (e.g., student work, free menstrual products, or voting age), based on focus group input (see Table A1 in Supplementary File 1 for an overview). De Jongerenstemtest also provided extra context and clarification for each policy statement, including explanations of the current situation for each policy, definitions of complex terms, and balanced pro/con arguments. This information was presented appealingly, with interactive swiping options and mimicking the aesthetic of VRT's Instagram-based news page NWS NWS NWS (381,000+ followers), featuring dynamic visuals, vibrant colors, and additional elements of interactivity. This approach aligns with the news consumption habits of young people, 87.1% of whom report daily Instagram use, with NWS NWS NWS being the most cited source (Jacobs et al., 2024).

In contrast, De Stemtest also provided information about the statements, which were presented in plain text in a toggle below each statement and employed a more restrained, less interactive, and less colorful design. Furthermore, De Stemtest allowed users to give more weight to issues, but the more complex feature was omitted from De Jongerenstemtest. Figure A1 in the Supplementary File 1 provides a visual comparison between the two VAAs.

3.3. *Dependent Variable: PPK*

While some studies measure PPK through participants' subjective feelings about how up to date they believe they are with parties' positions (Kamoen et al., 2015), we use a validated approach to measure PPK (Heinsohn et al., 2019; Munzert et al., 2021; Schultze, 2014). We asked respondents to indicate whether they believe the seven largest parties in Flanders support or oppose six specific policy proposals. These parties include the radical left PvdA, the green party Groen, the socialist party Vooruit, the Christian-Democratic party CD&V, the liberal party Open Vld, the nationalist and conservative party N-VA, and the radical right party Vlaams Belang.

One example of such a policy proposal is: "People with large assets should pay more taxes." An overview of the six statements is provided in Table A2 of Supplementary File 1, which includes additional features of the statements. Respondents had to indicate for each party whether it agreed or disagreed with the statement; a "don't know" option was available. Every respondent provided 42 answers, as participants had to estimate the party position for six statements related to seven parties. Statements were selected to strike a balance across policy issues (Van Camp et al., 2014); half of the statements focused on socio-economic issues (such as taxation, labor policies, and workers' rights) and the other half on socio-cultural matters (including climate policies, migration, and the legalization of soft drugs). Half of the statements were phrased so that agreeing would indicate a right-wing stance, while the other half were framed such that agreeing would imply a left-wing position. Important for our experimental set-up, three of the six statements featured in De Jongerenstemtest, and the other three were included in De Stemtest.

Participants' answers were then compared to the party positions. The party leaders provided these party positions, and hence, these are their official stances (there was cross-validation with the party manifesto). See Table A3 in Supplementary File 1 for an overview of these answers by party. The comparison between answers and the correct answer resulted in a binary variable, where 0 indicated incorrect answers ("wrong" or "don't know") and 1 indicated correct answers. The dataset was hierarchically structured by stacking it, resulting in a long format where each respondent received a score for every party on each policy statement.

3.4. *Independent Variables*

For *party type*, we created a dummy variable that classifies the three traditional parties—CD&V, Vooruit, and Open Vld—as mainstream, and PvdA, Groen, N-VA, and Vlaams Belang as niche. This classification follows prior work (Delwit & Van Haute, 2021; Luypaert, 2025) and reflects the distinction between parties embedded in traditional societal cleavages (Lipset & Rokkan, 1967) and those emphasizing narrower, issue-centered agendas. Mainstream parties have historically dominated government coalitions and pursued broad office- and policy-seeking strategies, whereas niche parties mobilize around unambiguous, often non-economic issues that are less prioritized by mainstream actors (Guth & Nelsen, 2021; Stiers, 2025). We acknowledge that this boundary is not always clear-cut and that parties such as Groen and N-VA occupy

more hybrid positions. Still, consistent with dynamic and saliency-based understandings of nicheness (Bischof, 2017; Meyer & Miller, 2015; Wagner, 2012), both can be considered niche. Ecological and regional segments constitute classic arenas of niche competition (Bischof, 2017): Groen focuses on environmental and post-materialist issues along the sociocultural GAL–TAN axis, whereas N-VA retains an ethno-regional and identity-based profile emphasizing Flemish autonomy, typically framing other issues—including socioeconomic ones—from a communitarian perspective (Abts et al., 2019). However, both have held government office, niche concerns, and programmatic focus rather than office exclusion (Meyer & Miller, 2015).

Party extremity is operationalized as the ideological distance between a party's mean position and the center of the left–right spectrum. Party positions were derived from *De Stemming 2025*, based on the average score of each party across 30 policy statements—15 related to socioeconomic issues and 15 to sociocultural topics—measured on a scale from –5 (left/libertarian) to +5 (right/authoritarian; Lefevere et al., 2025). Again, official party stances provided by the party leaders were used. Extremity was calculated as the absolute distance to the ideological center (0), resulting in two continuous variables: socioeconomic extremity and sociocultural extremity. See Table A4 and Figure A2 in Supplementary File 1.

A third independent variable indicates whether a *PPK knowledge item* was derived from *De Stemtest* or *De Jongerenstemtest*.

A final independent variable indicates whether a *policy statement* referred to the socioeconomic or sociocultural issue dimension, corresponding to the classical economic left–right divide and the GAL–TAN dimension (Kurella & Rapp, 2025). The socioeconomic dimension encompassed policies related to taxation, as well as labor regulations and workers' rights. The sociocultural dimension included policy statements on climate change and pollution, migration, and the legalization of soft drug use (see Table A2 in Supplementary File 1).

3.5. Randomization and Manipulation Check

Since we drew our respondents from a youth panel recruited through a convenience sample, our data are not representative of the broader youth population; highly educated and politically interested youngsters are overrepresented. Although this selective composition limits external validity, it also represents a least-likely case for detecting VAA-induced learning effects, as a high baseline of political interest reduces the room for improvement, making this a conservative test for finding VAA effects. Furthermore, non-parametric Kruskal–Wallis tests and chi-square tests indicate no significant differences between our experimental and control groups on key sociodemographic and attitudinal characteristics (see Table B1 in Supplementary File 2). Respondents in the three groups do not differ significantly in terms of age, gender, migration background (having at least one parent or grandparent born outside Belgium), mother's education level (highest diploma attained), financial stress (measured on an 11-point scale from *very difficult* to *very easy*), political interest (measured on an 11-point scale from *not at all interested* to *very interested*), or left–right self-placement on both the socioeconomic and sociocultural axis (ranging from 0 to 10). This indicates that our randomization was successful, mitigating concerns about the limitations of using a convenience sample; differences in PPK are likely to be attributed only to the treatment. Moreover, a recent study (Krupnikov et al., 2021) has verified that in political science research, convenience samples act similarly to representative samples. Given the absence of significant abnormalities in the randomization process, we report results without covariates.

We also conducted a manipulation check to determine whether respondents recognized that the VAA they participated in was aimed at the general or youth population. To this end, we employed two different manipulation checks. First, we asked respondents three questions, which they answered after participating in the VAA. The answers were given on an eleven-point scale ranging from 0 (*fully disagree*) to 10 (*fully agree*). The statements were the following: (a) “The VAA statements dealt with issues that youngsters consider important”; (b) “The VAA statements are comprehensible for youngsters”; and (c) “The VAA is visually attractive to youngsters.” Regarding these three items, respondents in the De Jongerenstemtest condition reported higher mean scores than those in the De Stemtest condition. ANOVAs revealed that these differences between the experimental groups were statistically significant ($F(1, 949) = 81.94, p < .001$; $F(1, 949) = 93.55, p < .001$; $F(1, 949) = 133.66, p < .001$, respectively). The data are visually reported in Figure 2.

Second, we also asked a more straightforward question to assess whether the manipulation worked: “Please indicate on the scale below to what extent the VAA was aimed at young people or adults,” with a score of 0 indicating that the VAA was aimed at youngsters and a score of 10 indicating it was aimed at adults. Here too, as can be seen in Figure 2, respondents in the De Jongerenstemtest condition perceived the VAA as more youth-oriented. In contrast, those in the De Stemtest condition viewed it as more adult-oriented (significant difference: $F(1, 949) = 218.03, p < .001$). Taken together, these results confirm that participants

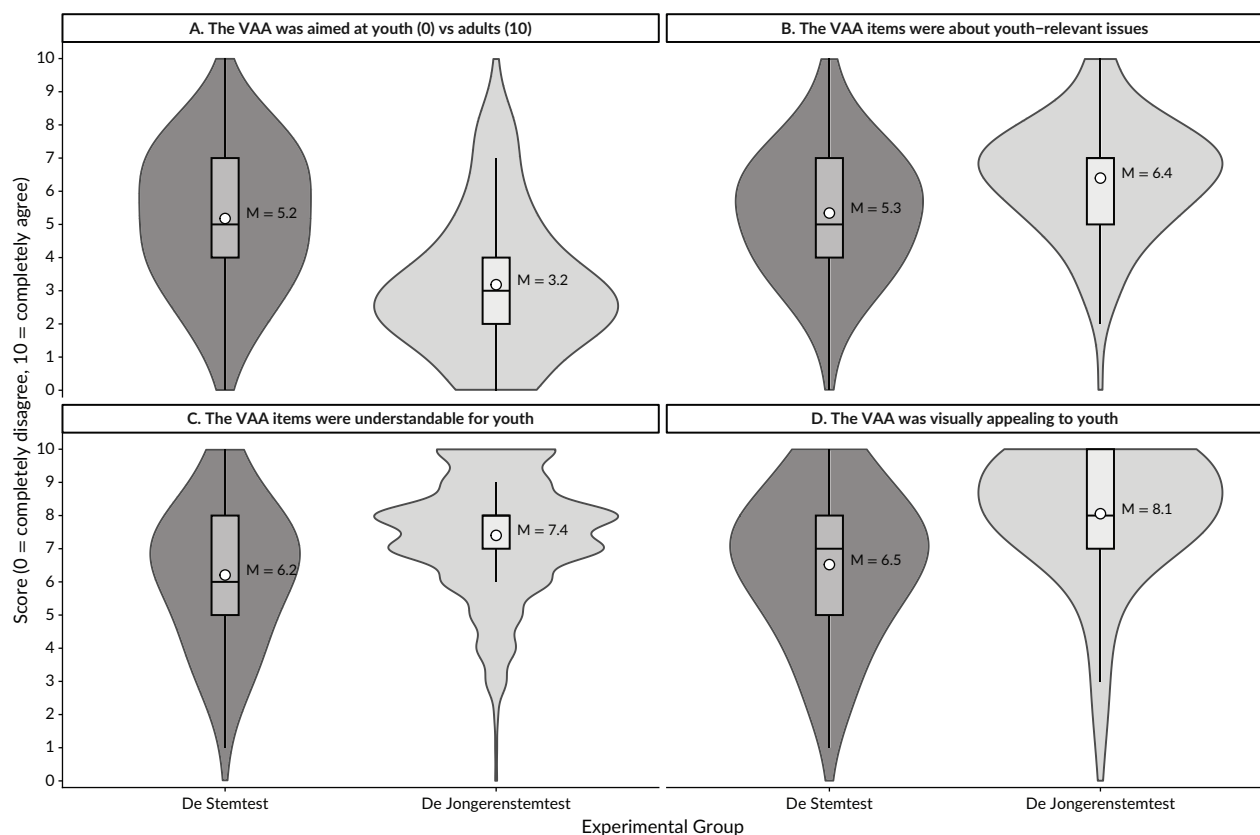


Figure 2. Manipulation check results for the experimental groups ($N = 951$): Panel A: general item (VAA aimed at youth vs adults); Panel B: youth-relevant issues; Panel C: understandable for youth; Panel D: visually appealing to youth. Note: Violin plots display the distribution of responses (boxplots = median/IQR; white circles = means).

accurately identified the target audience of the VAA they engaged with, demonstrating that our manipulation was successful.

3.6. Analyses

In Section 4, we analyze item–party responses using multilevel logistic regression with a random intercept for respondents (q_p_id) to account for repeated measures within individuals. We report odds ratios with 95% confidence intervals. The dependent variable is a binary indicator of PPK (1 = correct attribution; 0 = incorrect or “don’t know”); for interpretation, we present model-based predicted probabilities with 95% CIs. First, we describe baseline PPK patterns in the control group and estimate baseline models on the control sample only to assess variation by issue type, party, ideological extremity (economic and sociocultural), and party type (niche vs. mainstream). We then estimate experimental models comparing De Stemtest and De Jongerenstemtest to the control group, including interactions with whether the knowledge item was included in the assigned VAA, as well as probe moderation by party characteristics. Model fit and variance inflation factors are evaluated and reported in Supplementary File 2 (see tables B2, B3, B4, B5, B6, and B7); the null models of the included tables are reported in Table B8 in Supplementary File 2 and show substantial between-respondent variance ($ICC \approx .21-.22$), justifying the use of multilevel logistic regression models with random intercepts; statistical significance is evaluated at $\alpha = .05$ (two-tailed). Finally, analyses use listwise deletion.

4. Results

4.1. Non-Experimental Party and Issue Differences in PPK

To establish a baseline for evaluating the influence of VAA exposure, we first look at PPK among respondents in the control group. This subgroup’s PPK reflects participants’ unaided ability to identify party positions. The descriptive analysis reveals very large differences across political parties and issues (see Figure 3).

Regarding some socioeconomic issues, such as wealth taxation, correct attribution rates were high for PVDA (75%), N-VA (69%), and Vlaams Belang (62%), but low for mainstream parties, including CD&V (28%) and Open VLD (18%). Conversely, for student work flexibility, correct identification was poor across all parties, with correct responses ranging from 19% (PVDA) to 41% (Open VLD), and “don’t know” responses exceeding 50% for most parties. Similarly, performance on sociocultural issues was mixed. On the EU migration item, Vlaams Belang’s position was correctly identified by 85% of respondents, while Vooruit’s position was only correctly identified by 8%, with 64% giving incorrect responses. For items related to nitrogen emissions and cannabis legalization, correct attribution rates were generally much lower. For instance, correct identification of cannabis legalization ranged from 17% (Open VLD) to 43% (N-VA, Vlaams Belang), with “don’t know” responses exceeding 50% for many parties. Across various issues, parties at the ideological extremes (PVDA, Vlaams Belang, Groen) tended to be more easily recognized, particularly on topics such as wealth taxation and EU migration. In contrast, mainstream parties (CD&V, Open VLD, Vooruit) were more often met with uncertainty, as evidenced by higher “don’t know” rates. Although we cannot conduct a stringent test due to the limited number of issues, evidence suggests that respondents are more knowledgeable about a party’s position when that party “owns” the specific issue. These descriptive results highlight considerable variation in PPK depending on the issue and party.

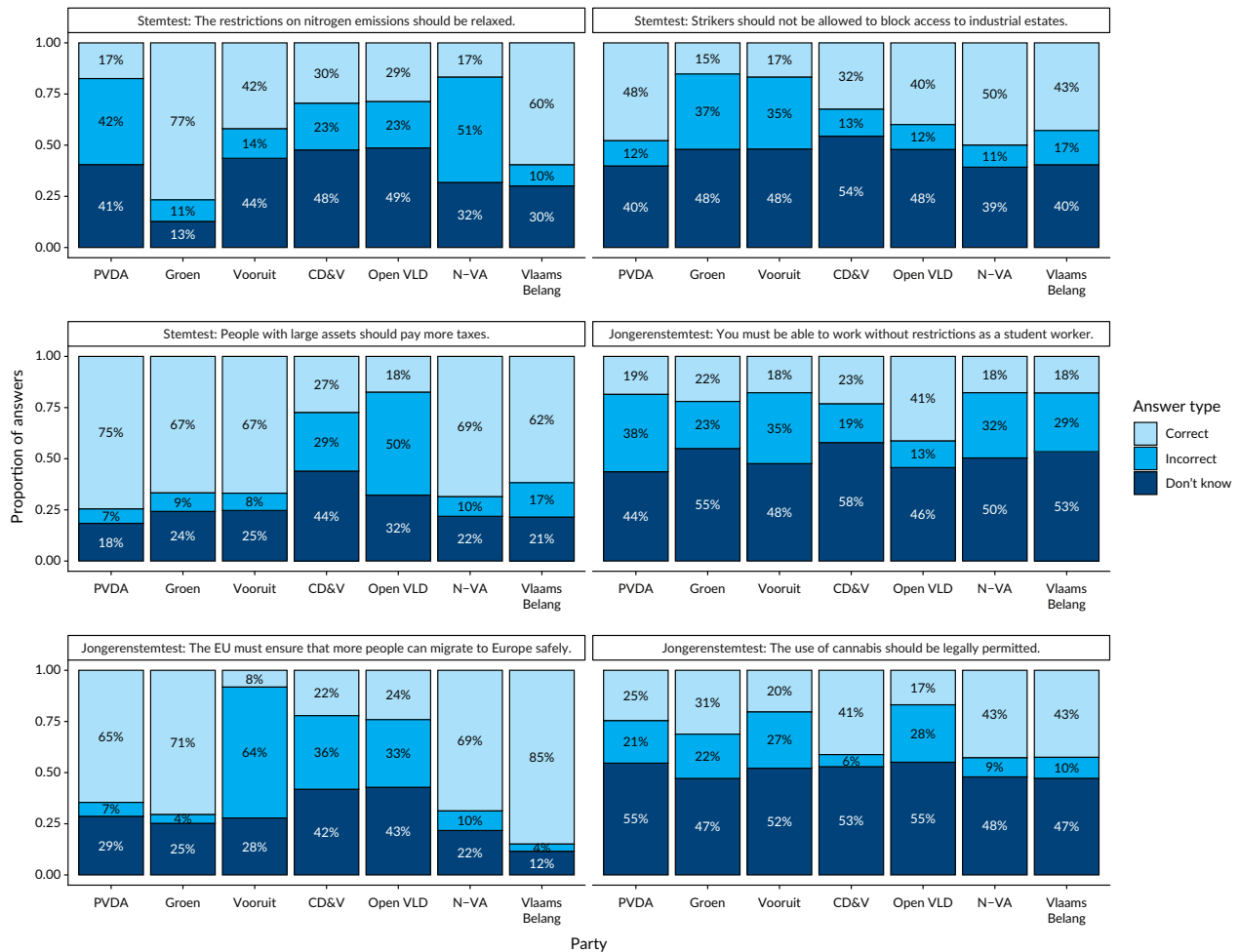


Figure 3. Proportion of correct, incorrect, and don't know answers by statement and party in the control group ($N = 1340$). Notes: Bars show the distribution of responses to each party's position, "correct" matches the party's official position, "incorrect" reflects a wrong attribution, "don't know" indicates the explicit don't know option.

Building on these descriptive patterns, we estimated multilevel logistic regression models to explain variation in PPK across issue and party characteristics. The main results are presented in Table B9 in Supplementary File 2. Respondents were less likely to correctly identify party positions on items from De Jongerenstemtest than on items from De Stemtest, showing 42% lower odds of correct attribution ($OR = 0.58, p < .001$; Model 1, Table B9 in Supplementary File 2). Knowledge was also somewhat higher for sociocultural items compared to socioeconomic ones ($OR = 1.34, p < .001$), suggesting that culturally oriented statements were more easily recognized. Positions of Groen ($OR = 1.33, p < .001$), N-VA ($OR = 1.15, p < .001$), and Vlaams Belang ($OR = 1.66, p < .001$) were more often correctly identified than those of PVDA. In contrast, mainstream parties such as Vooruit ($OR = 0.52, p < .001$), CD&V ($OR = 0.54, p < .001$), and Open VLD ($OR = 0.50, p < .001$) were less easily recognized. These party results are also shown in Figure 4, Panel A.

Party extremity was positively associated with knowledge. Model 2 shows that respondents were more likely to correctly attribute positions of parties with extreme stances on both economic ($OR = 1.14, p < .001$) and cultural dimensions ($OR = 1.36, p < .001$). These relationships were further refined in Model 3 and are shown in Figure 4, panels C and D, which indicates issue-specific effects: economic extremity was most predictive for

socioeconomic items ($OR = 0.78, p < .001$), whereas cultural extremity was most predictive for sociocultural items ($OR = 1.45, p < .001$). Model 4 depicted in Figure 4, panel B, demonstrates that niche parties were more easily recognized than mainstream ones, with more than twice the odds of correct attribution ($OR = 2.43, p < .001$).

These findings provide strong evidence for systematic variation in PPK depending on party and issue characteristics. Across all four models, the positions of parties that can be classified as niche, ideologically coherent, and positioned at the extremes of the political spectrum were more easily identified than those of

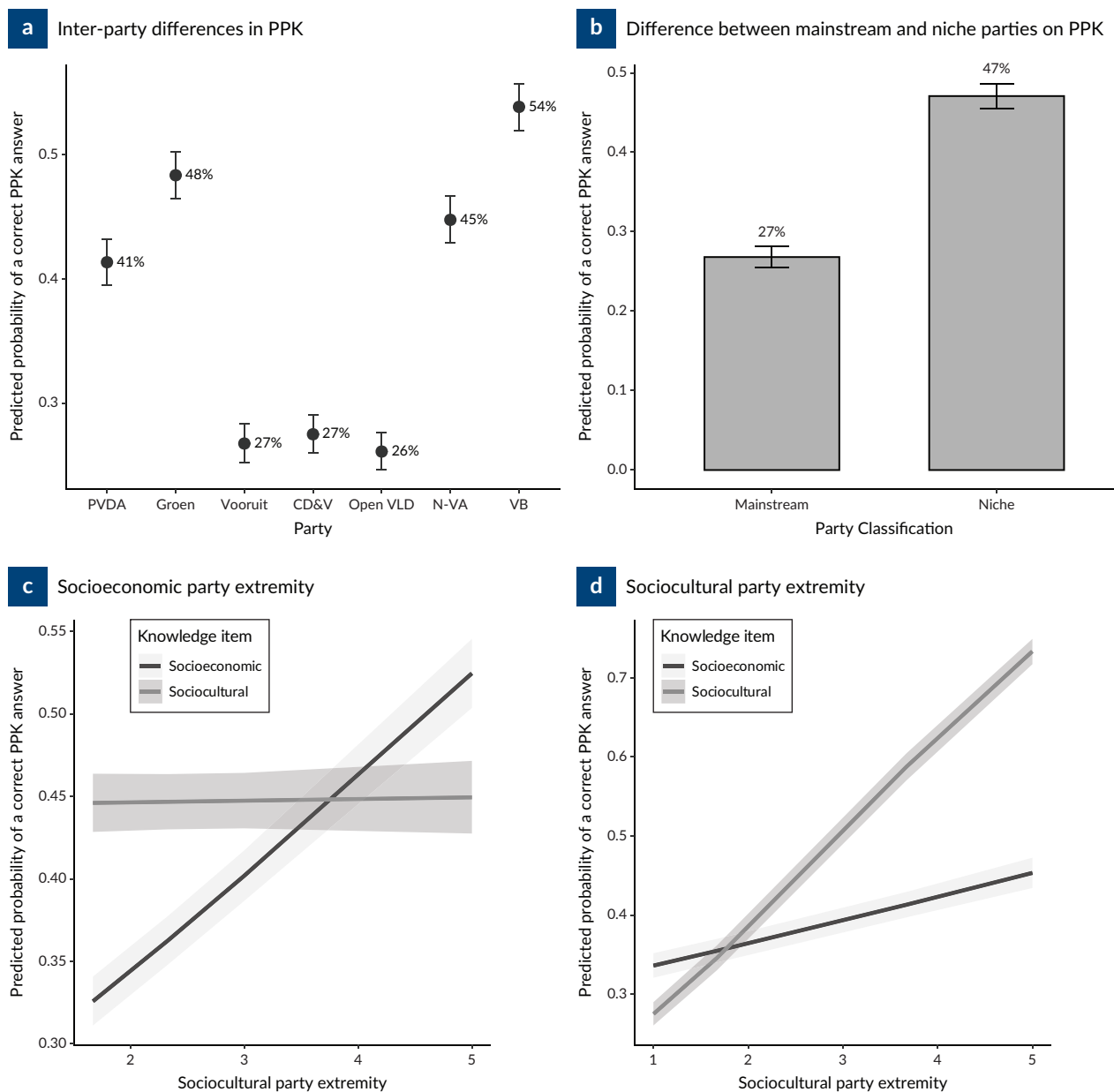


Figure 4. Predicted probabilities of PPK in control group participants ($N = 1340$): Panel A: inter-party differences; Panel B: mainstream versus niche parties; Panel C: socioeconomic party extremity by knowledge item; Panel D: sociocultural party extremity by knowledge item. Note: Estimates are based on mixed-effects logistic regression models (see Table B9 in Supplementary File 2).

mainstream and centrist parties. To assess whether our results depend on classifying Groen and N-VA as niche parties, we re-estimated all models treating both parties as mainstream. The substantive patterns remain virtually identical—effects shift slightly in magnitude but not in direction or significance (see Table B10 in Supplementary File 2). This confirms that the findings are robust to alternative operationalizations of “niceness.” Similarly, party extremity on both economic and cultural dimensions was positively associated with knowledge, particularly when matched with corresponding issue types. Because the nitrogen emissions statement could also be considered socioeconomic rather than sociocultural, we re-estimated all models with this item reclassified accordingly (see Table B11 in Supplementary File 2). Across all model specifications, the odds ratios for the sociocultural–socioeconomic contrast increase (e.g., from $OR = 1.34$ to $OR = 2.70$ in Model 1). Importantly, all substantive conclusions remain unchanged: knowledge of party positions is consistently higher for niche and ideologically extreme parties than for mainstream parties, regardless of how the nitrogen item is coded. Taken together, these results support Hypothesis 1, as PPK is indeed higher for niche political parties compared to mainstream parties.

4.2. Experimental Effect of VAA Use on PPK

We now turn to the experimental findings comparing the PPK of the control and treatment groups. This section tests Hypothesis 2, which held that VAA exposure increases PPK, Hypothesis 3, which posited that a youth-oriented VAA (De Jongerenstemtest) would have a stronger effect on youth than a generic VAA (De Stemtest), and Hypothesis 4, claiming that VAA learning effects would be moderated by party type (mainstream vs. niche).

Model 1 of Table B12 in Supplementary File 2 examines the overall effect of VAA exposure on PPK, regardless of whether the knowledge items were included in the VAA presented to respondents. It tests for potential learning effects that extend beyond the specific content of the VAA itself. Results show no significant difference between the group exposed to the generic VAA (De Stemtest) and the control group ($OR = 0.99$, $p = .895$). But, in contrast, respondents in the youth VAA group (De Jongerenstemtest) demonstrated a modest but significant increase in correct responses compared to the control group ($OR = 1.14$, $p = .020$). These findings imply that Hypothesis 2 is only partially supported, as an overall effect of VAA exposure on PPK was observed only for the youth VAA. Panel A of Figure 5 visualizes this overall treatment effect: the predicted probability is nearly identical for the control group and De Stemtest, while De Jongerenstemtest shows a slightly higher probability (approximately 30%).

Model 2 extends the analysis by adding an interaction between treatment group and knowledge item type. This interaction was included to examine whether respondents improved particularly on the items that were included in their assigned VAA, under the assumption that exposure effects would be strongest for these items. For De Stemtest, again, no significant interaction was found ($OR = 0.94$, $p = .110$), suggesting no additional benefit for items included in the generic VAA. However, a significant interaction was observed for De Jongerenstemtest ($OR = 1.14$, $p < .001$), indicating that participants in the youth VAA group were better able to identify party positions on the items they had been confronted with during the VAA. It is essential to note that the baseline accuracy for De Stemtest items was already higher in the control group, leaving less room for improvement compared to the De Jongerenstemtest items, which started at a lower level. These findings provide support for Hypothesis 3, demonstrating that the youth-oriented VAA was more effective in improving knowledge, particularly for its own included items. These item-specific treatment effects are

visualized in Panel B of Figure 5, where the predicted probabilities are separated by item type: general VAA items (circles) consistently show higher baseline accuracy. In contrast, youth-VAA items (triangles) start lower but increase notably in the De Jongerenstemtest condition.

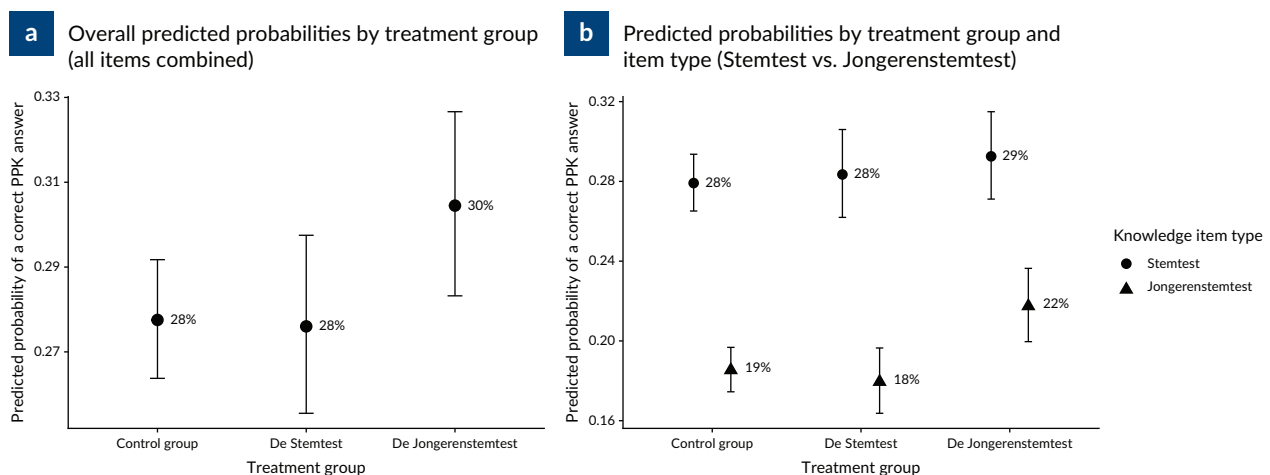


Figure 5. Predicted probabilities of correctly identifying party positions ($N = 2291$): Panel A displays overall treatment group differences, Panel B displays the treatment effects separately for general VAA items (*De Stemtest*, circles) and youth VAA items (*De Jongerenstemtest*, triangles). Note: Estimates are based on mixed-effects logistic regression models (see Table B12 in Supplementary File 2).

We now move on to testing Hypothesis 4 and further examining the potentially moderating role of party characteristics. As earlier results demonstrated that only exposure to *De Jongerenstemtest* was associated with improved PPK, these additional analyses focus exclusively on the youth VAA treatment group compared to the control group. The full output of the models we ran to test this hypothesis is presented in Table B13 in Supplementary File 2. Here, we only show the relevant figures.

In Figure 6, we present the results of a model that includes interaction terms between the treatment group and each party. None of these interactions reached statistical significance (all $p > 0.05$), suggesting that the knowledge gains associated with the *De Jongerenstemtest* were consistent across parties. As a robustness check, we re-estimated the interaction model using CD&V as the reference category. As reported in Table B14 in Supplementary File 2, this alternative specification yields the same substantive conclusion: none of the treatment and party interactions reach statistical significance.

Finally, Figure 7 underneath examines whether the learning effect differed between mainstream and niche parties. The interaction between exposure to *De Jongerenstemtest* and niche party classification was non-significant ($OR = 0.97$, $p = .594$). This suggests that the treatment effect was uniform across party types. Overall, these results do not support Hypothesis 4. The learning effect of *De Jongerenstemtest* appears to reflect a general improvement in PPK rather than targeted effects for specific parties or party types.

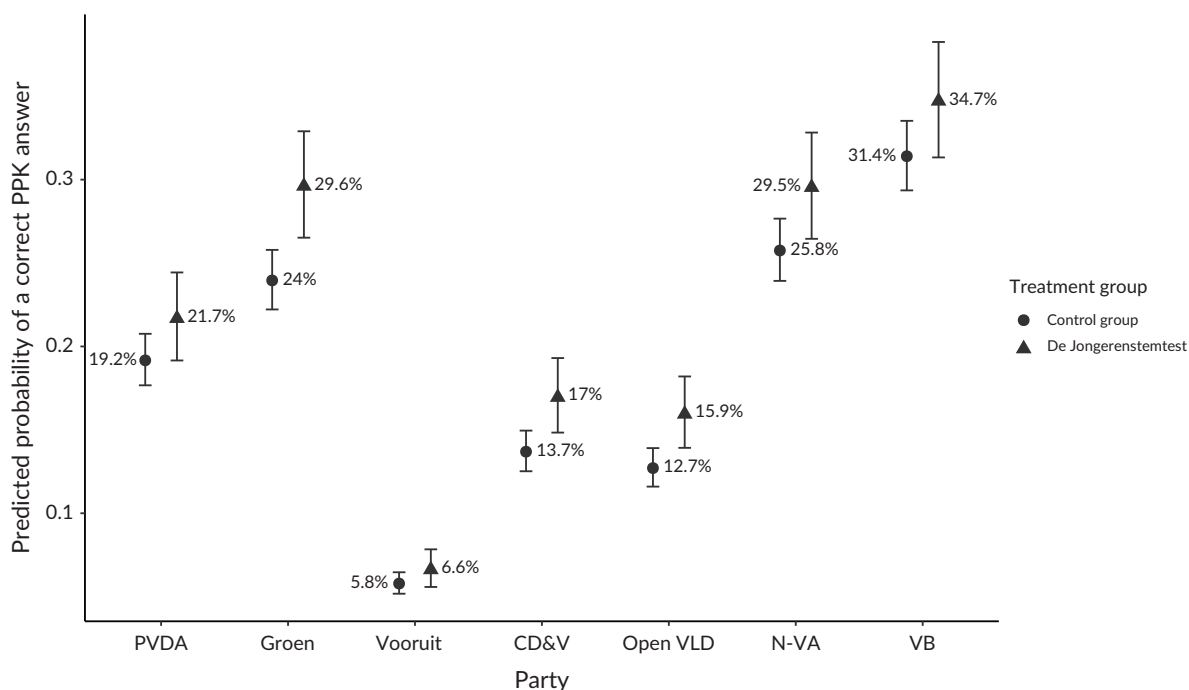


Figure 6. Predicted probabilities of PPK by party and treatment group ($N = 1827$). Notes: Estimates are based on mixed-effects logistic regression (see Table B13 in Supplementary File 2), circles represent the control group, and triangles represent the De Jongerenstemtest group, percentages indicate predicted probabilities of providing a correct answer for each party.

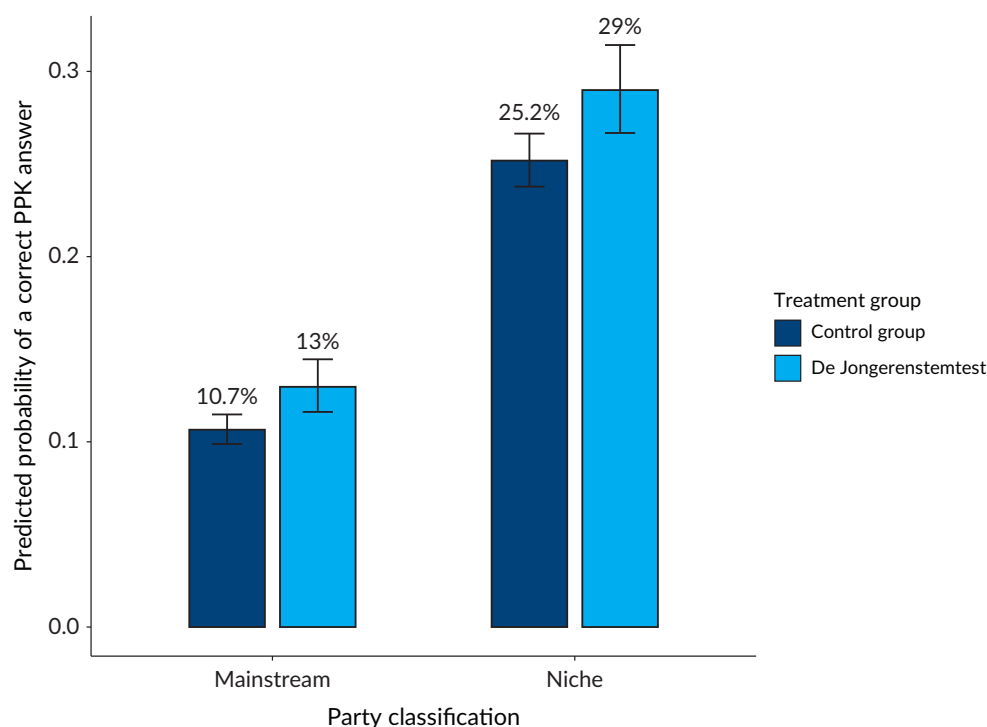


Figure 7. Predicted probabilities of PPK by party classification and treatment group ($N = 1827$). Notes: Estimates are based on mixed-effects logistic regression (see Table B13 in Supplementary File 2), dark bars represent the control group, while light bars represent the De Jongerenstemtest group, percentages above each bar indicate predicted probabilities of providing a correct answer.

5. Conclusion

PPK is essential for representative democracy to function properly. It allows voters to make a party choice that matches their policy preferences, thereby increasing the likelihood that the actual policies passed by parties align with popular opinion. Most work on PPK has focused on the individual drivers of PPK, but some studies have also examined the information that voters are exposed to. This study examined one type of information—yet an increasingly popular and, in some countries, even dominant type—VAAs. Millions of people use these applications when elections are due, but do these online algorithms make people better informed about party positions?

Our study shows that they do, under certain circumstances. Our experimental design with young Belgian voters revealed that those who were exposed to a youth VAA did indeed see their PPK increase. However, the young voters who used a generic VAA not specifically targeted at the young did not exhibit any increase in PPK. This suggests that the degree of engagement that a VAA can elicit may be key to explaining how PPK learning through VAAs works. VAAs that manage to draw attention and appear to be relevant are better at generating knowledge effects. Our experimental design cannot tell us what exactly is in the youth VAA we used that causes the effect. The youth VAA we used presented a bundled treatment that differed in quite some aspects from the generic VAA. The selected policies were more relevant for youngsters, more information was given, the VAA was simpler, the layout and look were different, and there were more interactive options. Follow-up work should aim to identify what aspects of VAAs capture people's attention and engage them, as this appears to be crucial for explaining learning through VAA use. A further limitation concerns the durability of these learning effects: as our design captures outcomes immediately after VAA use, we cannot assess how long such gains persist. Future work should incorporate follow-up measurements to examine their longevity.

Another unknown is how the learning through VAAs exactly works. Do people learn indirectly through the revealed distance between their own positions and by attributing their own positions to the parties they were told to be close to (and vice versa)? Or is it only those who dug deeper and used the additional features of direct information about parties' positions who learned a few things? We cannot tell based on our findings. However, we hope future work will take up the challenge to identify the exact mechanisms. Future research should examine more closely why citizens seem more knowledgeable about niche parties than mainstream parties, for example, by investigating how young people inform themselves during election campaigns and the specific media diet and sources they rely on. In this article, we provided only a general test based on six issues, but our results also cautiously suggest that—while overall young citizens had more knowledge of niche parties—the clarity of party cues partly depends on issue ownership. Expanding the issue scope in future studies could help assess these patterns more systematically, providing a more nuanced understanding (see Rovny & Polk, 2020).

Our study dealt with young voters. While there are good reasons to focus on youngsters, the question remains to what extent our results can be generalized to voters more generally. Again, we cannot tell. We did not have adult subjects and cannot tell whether they would have increased their PPK after being exposed to the generic (or the youth) VAA. However, we argue that the logic of engagement would also likely apply to adult voters. The more a VAA, through its content and design, manages to engage its users, the more likely it is that positive effects on knowledge will emerge.

Acknowledgments

The authors would like to thank the entire YouVoice team at the University of Antwerp and KU Leuven for their valuable feedback and support, in particular, Cato Waeterloos (KU Leuven), Margot Lissens (KU Leuven), Babette Hermans (KU Leuven), Michaël Opgenhaffen (KU Leuven), Desiree Schmuck (University of Vienna), and Peter Van Aelst (University of Antwerp). We would also like to thank Jonas Lefevere (University of Antwerp) and the Bpact team for their valuable support while collecting our experimental data. We also thank the entire Stemtest team that helped create De Stemtest and De Jongerenstemtest. Finally, we are grateful to the Voting Advice Applications Research Network for their constructive feedback on an earlier version of this article during the ECPR General Conference in Dublin (2024).

Funding

This research was supported by the inter-university iBOF program (grant number iBOF/23/036: FN5429), with Peter Van Aelst as the principal investigator. Publication of this article in open access was made possible through the institutional membership agreement between the Vrije Universiteit Brussel and Cogitatio Press.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability

Due to the nature of the research and in accordance with ethical and legal restrictions, supporting data are not available. We report data on minors (youngsters from 16 to 30), but we are restricted in sharing our data due to the sensitive nature and restrictions imposed by our ethical board and university privacy commission.

LLMs Disclosure

We used OpenAI's ChatGPT (GPT-5) as a supportive tool to streamline R code for data preparation and analysis; the authors made all substantive decisions and interpretations.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Abts, K., Dalle Mulle, E., & Laermans, R. (2019). Beyond issue diversification: N-VA and the communitarianisation of political, economic and cultural conflicts in Belgium. *West European Politics*, 42(4), 848–872.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.
- Barabas, J., & Jerit, J. (2009). Estimating the causal effects of media coverage on policy-specific knowledge. *American Journal of Political Science*, 53(1), 73–89.
- Barabas, J., Jerit, J., Pollock, W., & Rainey, C. (2014). The question(s) of political knowledge. *American Political Science Review*, 108(4), 840–855.
- Bischof, D. (2017). Towards a renewal of the niche party concept: Parties, market shares and condensed offers. *Party Politics*, 23(3), 220–235.
- Bischof, D., & Senninger, R. (2018). Simple politics for the people? Complexity in campaign messages and political knowledge. *European Journal of Political Research*, 57(2), 473–495.
- Boonen, J. (2015). The development of stable party preferences: Explaining individual-level stability among adolescents in Belgium. *YOUNG*, 23(4), 313–335.

- Conover, P. J., & Feldman, S. (1989). Candidate perception in an ambiguous world: Campaigns, cues, and inference processes. *American Journal of Political Science*, 33(4), 912–940.
- Dahlberg, S., & Hartevelde, E. (2016). Left-right ideology as an inferential device in multiparty systems: Can citizens overcome low information by imputing parties' policy positions? *Electoral Studies*, 42, 175–187.
- Dalton, R. J., & McAllister, I. (2015). Random walk or planned excursion? Continuity and change in the left-right positions of political parties. *Comparative Political Studies*, 48(6), 759–787.
- DeCamp, W., & Manierre, M. J. (2016). "Money will solve the problem": Testing the effectiveness of conditional incentives for online surveys. *Survey Practice*, 9(1), 1–9.
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Delwit, P., & Van Haute, E. (2021). *Les partis politiques en Belgique*. ULB.
- Dennison, J., & Kriesi, H. (2023). Explaining Europe's transformed electoral landscape: Structure, salience, and agendas. *European Political Science Review*, 15(4), 483–501.
- Deschouwer, K. (2012). *The politics of Belgium: Governing a divided society*. Palgrave Macmillan.
- Enyedi, Z. (2016). The influence of voting advice applications on preferences, loyalties and turnout: An experimental study. *Political Studies*, 64(4), 1000–1015.
- Garry, J., Tilley, J., Matthews, N., Mendez, F., & Wheatley, J. (2019). Does receiving advice from voter advice applications (VAAs) affect public opinion in deeply divided societies? Evidence from a field experiment in Northern Ireland. *Party Politics*, 25(6), 854–861.
- Garzia, D., & Marschall, S. (2016). Research on voting advice applications: State of the art and future directions: Research on voting advice applications. *Policy & Internet*, 8(4), 376–390.
- Guth, J. L., & Nelsen, B. F. (2021). Party choice in Europe: Social cleavages and the rise of populist parties. *Party Politics*, 27(3), 453–464.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014, January 6–9). *Does gamification work?—A literature review of empirical studies on gamification* [Conference paper]. Hawaii International Conference on System Sciences, Waikoloa, HI, United States.
- Heinsohn, T., Fatke, M., Israel, J., Marschall, S., & Schultze, M. (2019). Effects of voting advice applications during election campaigns. Evidence from a panel study at the 2014 European elections. *Journal of Information Technology & Politics*, 16(3), 250–264.
- Henderson, M. (2014). Issue publics, campaigns, and political knowledge. *Political Behavior*, 36(3), 631–657.
- Jacobs, L., Matthieu, J., Waeterloos, C., Lissens, M., Hermans, B., Van Campenhout, M., Opgenhaffen, M., Schmuck, D., Walgrave, S., & Van Aelst, P. (2024). *YouVoice / POLKNOW youth panel technical report*. University of Antwerp, KU Leuven.
- Jennart, I. (2025). *To know it all? How party position knowledge varies between citizens, political parties, and policy issues. And why it matters*. Universiteit Antwerpen.
- Jennings, M. K., & Stoker, L. (2004). Social trust and civic engagement across time and generations. *Acta Politica*, 39(4), 342–379.
- 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.
- Kivunja, C. (2014). Theoretical perspectives of how digital natives learn. *International Journal of Higher Education*, 3(1), 94–109.
- Krosnick, J. A. (1990). Government policy and citizen passion: A study of issue publics in contemporary America. *Political Behavior*, 12(1), 59–92.
- Krupnikov, Y., Nam, H. H., & Style, H. (2021). Convenience samples in political science experiments. In

- J. N. Druckman & D. P. Green (Eds.), *Advances in experimental political science* (pp. 165–183). Cambridge University Press.
- Kurella, A.-S., & Rapp, M. (2025). Unfolding GAL-TAN: The multi-dimensional nature of public opinion in Western Europe. *West European Politics*. Advance online publication. <https://doi.org/10.1080/01402382.2025.2466117>
- Lefevere, J. (2024). Is there clarity in ambiguity? A literature review of party positional ambiguity. *Party Politics*, 30(2), 334–344.
- Lefevere, J., & Verwee, A. (2025). Ambiguity in the campaign, and in office: How parties' ambiguous policy statements affect party support. *International Journal of Public Opinion Research*, 37(3), Article edaf032.
- Lefevere, J., Walgrave, S., & Pilet, J.-B. (2025). *De Stemming/L'Enquête Nationale 2025* (Versie 2) [Dataset]. Social Sciences and Digital Humanities Archive—SODHA. <https://doi.org/10.34934/DVN/VAHHXZ>
- Lipset, S. M., & Rokkan, S. (1967). Cleavage structures, party systems, and voter alignments: An introduction. In S. M. Lipset & S. Rokkan (Eds.), *Party systems and voter alignments: Cross-national perspectives* (pp. 1–64). The Free Press.
- Luypaert, J. (2025). Navigating new realities: Explaining programmatic transitions of mainstream and niche parties. *International Political Science Review*, 46(1), 125–143.
- Meyer, T. M., & Miller, B. (2015). The niche party concept and its measurement. *Party Politics*, 21(2), 259–271.
- Munzert, S., Barberá, P., Guess, A., & Yang, J. (2021). Do online voter guides empower citizens? *Public Opinion Quarterly*, 84(3), 675–698.
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706. <https://doi.org/10.1080/10584609.2020.1843572>
- Neundorff, A., & Smets, K. (2017). Political socialization and the making of citizens. In *Oxford handbook topics in politics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199935307.013.98>
- Pianzola, J. (2014). Selection biases in voting advice application research. *Electoral Studies*, 36, 272–280.
- Rovny, J. (2013). Where do radical right parties stand? Position blurring in multidimensional competition. *European Political Science Review*, 5(1), 1–26.
- Rovny, J., & Polk, J. (2020). Still blurry? Economic salience, position and voting for radical right parties in Western Europe. *European Journal of Political Research*, 59(2), 248–268.
- Schultze, M. (2014). Effects of voting advice applications (VAAs) on political knowledge about party positions: Effects of VAAs on political knowledge about party positions. *Policy & Internet*, 6(1), 46–68.
- Singer, E., & Kulka, R. A. (2002). Paying respondents for survey participation. In M. Ver Ploeg, R. A. Moffit, & C. F. Citro (Eds.), *Studies of welfare populations: Data collection and research issues* (Vol. 4, pp. 105–128). National Academy Press.
- Stiers, D. (2025). Why do voters support niche parties? *Party Politics*, 31(5), 881–891.
- 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 & S. Marschall (Eds.), *Matching voters with parties and candidates: Voting advice applications in comparative perspective* (pp. 11–31). The 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.
- Wagner, M. (2012). When do parties emphasise extreme positions? How strategic incentives for policy differentiation influence issue importance. *European Journal of Political Research*, 51(1), 64–88.
- Walgrave, S., Rihoux, B., Goovaerts, I., Jacobs, L., Lefevere, J., Matthieu, J., Mongrain, P., Tsoulou-Malakoudi, A., Van Campenhout, M., Uyttendaele, L., Nuytemans, M., & Henkens, E. (2025). *Stemtest 2024 / Test Electoral 2024* [Dataset]. SODHA. <https://doi.org/10.34934/DVN/SOCUOL>

Zaller, J. R. (1992). *The nature and origins of mass opinions*. Cambridge University Press.

About the Authors



Joke Matthieu is a postdoctoral researcher at the Department of Political Science (M²P research group) at the University of Antwerp and at the Department of Political Science (DFUTURE research group) at Vrije Universiteit Brussel. Her main expertise lies in inequalities within political socialization processes.



Laura Jacobs is a postdoctoral researcher and guest professor at the Department of Political Science and M²P research group of the University of Antwerp. Her research interests include anti-immigrant attitudes, media effects, voting for the radical right, effects of the cordon sanitaire, elite discourse, and political communication.



Matthias Van Campenhout is a PhD researcher affiliated with the M²P research group within the Department of Political Science at the University of Antwerp. His research primarily focuses on the political knowledge of (young) citizens.



Stefaan Walgrave is a professor of political science at the University of Antwerp and head of the M²P research group. His research focuses on social movements, political communication, elections, and political representation.

ARTICLE

Open Access Journal 

Are Single-Party Voting Advice Applications Useful? Comparing Voter Preferences in the BSW-O-Mat With a Probability Sample

Jan Philipp Thomeczek 

Faculty of Economics and Social Sciences, University of Potsdam, Germany

Correspondence: Jan Philipp Thomeczek (jan.philipp.thomeczek@uni-potsdam.de)

Submitted: 29 August 2025 **Accepted:** 27 November 2025 **Published:** 8 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

As the voting advice application (VAA) market has grown increasingly competitive, developers have focused on innovation, including specialised VAAs that focus on specific topics. This article evaluates the analytical potential of a “single-party VAA” called BSW-O-Mat, a website developed to assess alignment with the newly founded German party Bündnis Sahra Wagenknecht (BSW), attracting over 50,000 participants. Single-party VAAs face compounded selection bias; beyond the typical overrepresentation of educated politically engaged users, they disproportionately attract voters interested in the focal party. Despite over 50,000 responses, questions remain about data validity for estimating voter preferences. To address the usefulness of VAA data for analysing the new party’s supporters, this study examines convergence validity by comparing the average marginal effects from logistic regression models based on VAA data with those from the 2025 German Longitudinal Election Study’s (GLES) probability sample. Despite substantial overrepresentation of potential BSW supporters in the VAA sample, convergence in the direction of effects was observed across datasets. Effect correlations reached $r = 0.65$ (increasing to $r = 0.87$ when excluding non-converging effects), though the VAA dataset tends to overestimate effect sizes. Notably, even analysing substantially underrepresented AfD voters (5.5% in VAA vs 19.8% in GLES) yielded results largely consistent with probability-sample models. These findings demonstrate that the VAA data can provide valid insights when appropriately weighted and interpreted in light of self-selection biases.

Keywords

BSW; convergence validity; non-probability sample; selection bias; VAAs

1. Introduction

Over the last two decades, the number of VAA users has grown steadily, now reaching potentially millions of voters (Garzia & Marschall, 2019). Initially developed for civic education, especially for first-time voters, VAAs also provide campaign-related data for political scientists. VAAs offer valuable data on candidates' and parties' positions (Ferreira da Silva et al., 2021) and user data on the same issues. Some VAAs, such as the Wahl-O-Mat, do not directly collect user data or offer follow-up surveys only to randomly selected users. Others use opt-ins for data collection, often in combination with additional questionnaires. Such additional questions often relate to political attitudes, voting intention, and socio-demographic variables. As user and party data address the same issues, they can be used to analyse party-voter congruence (Tromborg & Albertsen, 2023; M. Wagner & Ruusuvirta, 2012; Zhirnov et al., 2025).

VAA user data presents unique opportunities for exploring the voter's mindset. However, the opt-in nature of VAA samples is not unproblematic. VAA usage is linked to several self-selection biases, as male, highly educated, and politically interested users are generally overrepresented (Marschall, 2014; Pianzola, 2014). Accounting for these biases can help to make the data "more" representative, for example, with survey weights (Toshkov & Romeijn, 2021).

During the last decade, the "VAA market" has become increasingly crowded, and for many elections, even at the subnational level, multiple tools are available. This has resulted in numerous "specialised" VAAs, tools dedicated to specific topics and content. For example, in 2023, Dutch VAA company Kieskompas launched a Climate and Energy tool that matched users and parties on issues related to climate change, environmental politics, and energy (Vogels, 2023). Other tools calculate matches based on historical contexts: the Federal Archive of Germany has published a tool that allows users to explore their closeness to parties from the Weimar Republic (Bundesarchiv, 2018).

Although the number of such specialised tools has increased strongly over the last few years, the usefulness and associated biases of user data from specialised VAAs remain largely unstudied. After all, such datasets are affected not only by known biases related to interest in elections but also by biases inherent to the specialised content itself. To use the cited example, a politically highly interested user may use a tool for climate issues only if they are interested in this topic. This could lead to a situation where pro-environmentalists are extremely overrepresented in this tool's dataset. While specialised VAAs can provide valuable data on topics that neither classic VAAs nor traditional surveys adequately cover, biases in their use may strongly limit the potential of these datasets.

The following article addresses this research gap by comparing the convergence of the analytical results from the specialised VAA BSW-O-Mat with those from a probability sample. BSW stands for Bündnis Sahra Wagenknecht, a party formally founded in January 2024. Shortly after the party was officially founded, political scientists from the University of Potsdam developed the VAA tool BSW-O-Mat. It can be described as a single-party VAA, as users are confronted with items derived from BSW's first manifesto. This facilitates the analysis of party-voter congruence, since VAA users have positioned themselves to the party's own policies.

However, the self-selection bias inherent in such a tool is obvious: those more curious about the new party were more likely to use it. While the dataset offers valuable insights into more than 50,000 users, it may

nevertheless be of little value if it is strongly skewed towards (likely) BSW voters and lacks those who would not vote for the party. This article puts the empirical usefulness of such a dataset to a hard test by comparing the analytical results based on it with a probability sample. Although VAA research has suggested that voter policy preferences obtained from weighted VAA datasets are comparable to probability sample estimates (Toshkov & Romeijn, 2021), testing this with a specialised VAA dataset represents a stricter test, as an additional bias compounds standard VAA biases.

The article is structured as follows. The following second section will summarise the research on the extent to which VAA data is useful for analysing voter preferences. This will be followed by Section 3, which identifies variables for the convergence validity test by discussing research on BSW's voters and their voting potential. The fourth section introduces the research design, which compares the BSW-O-Mat dataset with the German Longitudinal Election Study's (GLES) rolling cross-sectional dataset, a probability sample collected for the 2025 Bundestag elections. Section 5 provides a detailed discussion of the processes of data cleaning, sampling, and weighting. The core analysis then compares the results for four different models, focusing on socio-demographics, policy preferences, populism, and vote recall in predicting BSW support, which were calculated using the two datasets. The conclusions summarise the results and discuss their implications for future research.

2. The Challenges of Using VAA Data for Voter Research

Ladner and Fivaz (2012, p. 178) describe VAAs as “issue-matching systems.” This summarises the basic functionality of VAAs, as they are designed to produce voting advice based on the similarities between their users and political parties and/or candidates on the most salient political issues (Garzia & Marschall, 2019). VAAs typically ask users to state their agreement with 20–35 policy-related statements. Once a user completes the task, which is technically a questionnaire with a series of Likert-scale questions, the VAA describes the similarity between their position and the positions of the major political actors.

Most VAAs also collect user data, especially projects led by researchers. Data collection is typically opt-in, though some VAAs also collect data under an opt-out principle or ask only a random sample of VAA users to provide additional information. As tools such as the German Wahl-O-Mat reach tens of millions of voters (Marschall, 2014), the advantages of VAA user data are obvious. Such large-*n* datasets allow for the study of even relatively small subgroups of voters, such as those of smaller parties. This is often problematic with probability samples. As Toshkov and Romeijn (2021, p. 2) demonstrate, even analyses of voters of 5%-parties (e.g., junior coalition partners) are difficult, with standard sample sizes of $n = 1,000$ respondents, as estimated policy preferences have large statistical uncertainty.

One of the main limitations of VAA data is that VAA users deviate from the general population in several respects. Research has shown that VAA usage is systematically linked to variables such as education, political interest, age, and gender (Hooghe & Teepe, 2007; Marschall, 2014; Pianzola, 2014). In addition, VAA samples are opt-in samples by design. Everyone who has access to the VAA link can use the tool, and VAAs are typically designed to reach as many voters as possible. This stands in stark contrast to classic probability samples, which aim to gather responses from typically 1,000 to 2,000 respondents who should closely resemble the general population.

The extent to which these biases pose a problem depends on the research question. Munzert and Ramirez-Ruiz (2021) have noted that reported VAA effects on turnout and voting choice strongly differ between observational and experimental VAA data. While the former tend to overreport such effects, the latter represent a setting in which VAA usage is randomised, i.e., no self-selection into using the tool can occur. In analyses based on experimental VAA samples, effects on turnout and voting choice are much smaller and often not statistically significant. However, as VAAs are open to the public, a typical VAA dataset is always observational.

Wheatley et al. (2014) highlight that in VAA tools, selection bias is strongly associated with the intention to participate in the respective elections. This is reflected in the share of non-voters among VAA users. Gemenis and Rosema (2014) report that, according to the Dutch election study, only 3% of Dutch VAA users were non-voters. To be fair, non-voters are generally difficult to reach, also with probability samples (see Table 1). Gemenis and Rosema report that even among non-VAA users, only 9.4% of the respondents were self-declared non-voters. There appears to be a general trend showing that VAA data may be less useful for studying aspects unrelated to elections, which extends to sociological research (Sheppard, 2018). Another study by Pianzola (2014) suggests modelling the selection process, for example, related to political interest. This underlines one of the most crucial aspects regarding the usefulness of VAA data. To assess data quality, reference data are needed, which are not always available. Data should be collected around the same time, and questions should, ideally, be identical in wording.

There are, however, studies showing that such biases are less important than assumed. Toshkov and Romeijn (2021) analysed the policy preferences of party voters in Germany and the Netherlands and concluded that positions modelled with VAA data are comparable to those from representative voter studies. However, they do underline that, especially in the Netherlands, policy preferences based on weighted VAA data are closer to those obtained from representative surveys. Thomeczek et al. (2025) demonstrate that dimensions obtained from unweighted EU-wide VAA data are comparable to those derived from the European Election Study's sample.

The overall question is related to the validity of VAA data, which addresses two aspects. First, there is the question of whether VAA datasets can identify the same substantial effects as probability samples, with effects pointing in consistent directions. The second aspect is related to effect sizes, namely, whether the magnitude of these effects aligns with analyses based on probability samples. If VAA data correctly identify which factors matter and in which direction, researchers can use them to understand voting behaviour mechanisms, even if effect magnitudes differ due to selection bias or measurement contexts. Conversely, if different variables are significant across datasets, the data may produce similar predictions through compensating for errors, but would not support explanatory inferences about potentially causal drivers of voting.

While research has focused on the data quality of classic VAA datasets, the potential of datasets from specialised VAAs is largely understudied. The goal of this article is to address this research gap by focusing on the user data of the single-party VAA BSW-O-Mat. The results will be compared to analyses based on a probability sample from the GLES. This provides a hard test: if convergence validity is achieved despite heightened selection bias, it demonstrates that specialised VAA data can be used for substantive voter analysis when appropriately weighted. Conversely, failure to converge would indicate that single-party VAAs' self-selection overwhelms demographic corrections, limiting their analytical utility to descriptive

purposes. The next section summarises the literature on BSW voters to identify potential effects that will be analysed with the single-party VAA dataset and the GLES probability sample.

3. Who Supports the BSW?

Although the BSW is a comparatively young party, founded in early 2024, several published articles have examined its potential or actual voters. This section discusses the most important factors that explain BSW's voting potential for the convergence validity test. The following literature review, therefore, focuses on factors that (a) are theoretically established BSW predictors and (b) can be operationalised in both datasets using available items.

S. Wagner et al. (2023) found higher BSW potential among East Germans, cultural conservatives, and voters with pro-market, anti-immigration attitudes. This analysis preceded the party's official founding, when Wagenknecht had long contemplated launching such a party. Furthermore, those who are less satisfied with the democratic output and those with a higher propensity to vote for the AfD and Die Linke are reportedly more likely to vote for such parties. The ideological profile of such a new party was of particular interest, as many expected it to have potential appeal among "left-authoritarian" voters, who had not had a political home in Germany so far (Steiner & Hillen, 2021).

Academic interest in the party has been growing since its official founding in January 2024 (Patton, 2025). Most of the findings by S. Wagner et al. (2023) have been confirmed in subsequent publications (Heckmann et al., 2025; Herold & Otteni, 2025; Hoffmann, 2025; Jankowski, 2024; Steiner & Hillen, 2025). However, some additional factors have been identified by researchers as explaining BSW's appeal. Firstly, foreign policy plays an important role for potential BSW voters. Over the course of the Ukraine war, Wagenknecht has developed distinct foreign policy positions. She is highly critical of NATO and seeks common ground with Russia, advocating for "peace" in Ukraine, even at the cost of Ukrainian territory. Therefore, NATO scepticism and rejection of delivering more weapons to Ukraine are strong predictors of BSW support (Heckmann et al., 2025; Hoffmann, 2025; Wurthmann & Wagner, 2024). Secondly, the BSW offers an electorally attractive alternative to those holding populist attitudes (Herold & Otteni, 2025; Hoffmann, 2025; Jankowski, 2024; Thomeczek, 2024a) and having less institutional trust (Baumann & Seikel, 2024). Finally, regarding the BSW's potential among AfD voters, the effect depends on how AfD voters are operationalised. Those who contemplate voting for the AfD show a stronger propensity towards the BSW (Heckmann et al., 2025; Jankowski, 2024). However, the BSW has actually not won over many voters who voted for the AfD at the previous national elections in 2021 (Infratest dimap, 2025b; Jankowski, 2024). High BSW potential is reported among former left voters, non-voters, and voters of minor parties (Jankowski, 2024; Thomeczek, 2024b).

Some socio-demographic factors are relevant in explaining BSW support (Bundeswahlleiterin, 2024, 2025). Accordingly, the BSW tends to be voted for more by women than by men, and by East Germans than by West Germans. A significant shift can be observed in terms of age. While the BSW achieved its highest vote share among the baby boomer generation in the European elections, it received its highest vote share among 18 to 44-year-olds in the Bundestag elections.

Based on this literature, the convergence validity test will examine six variable domains: (a) socio-demographics, (b) policy preferences, (c) populist attitudes, institutional trust, and democratic satisfaction, as well as (d) vote recall (2021 Bundestag election). The next section details how these are operationalised in both datasets.

4. Research Design

The research design follows Toshkov and Romeijn (2021) and tests convergence validity (Adcock & Collier, 2001) between the analysis based on the single-party VAA dataset and the probability sample. Convergence validity tests whether two different data sources produce similar results when analysing the same (political) phenomenon. If they converge, it is assumed that both measurements are valid. For the following analysis, several factors impacting the BSW vote with matching variables in both datasets are analysed, and their results are compared. The extent to which they arrive at similar conclusions is then related to the extent to which we can speak of convergence validity.

In this study, two aspects of convergence validity are examined to explain BSW's voting potential: substantial effects and effect sizes. Given the exploratory nature of this validation—the first for a single-party VAA—we refrain from specifying ex-ante thresholds for “acceptable” convergence. Instead, we report convergence descriptively and interpret patterns in light of theoretical expectations about selection bias. Substantial effect convergence is related to the identification of a statistically significant effect: do the BSW-O-Mat and GLES datasets yield comparable findings regarding the identification of factors significantly affecting voting for BSW? If they do, this suggests that both datasets capture genuine voter preferences rather than measurement artefacts. The second part of the analysis focuses on the convergence of effect sizes. As a last step, it will also be tested whether the dataset can achieve convergence validity when analysing AfD voters, a strongly underrepresented voter group in the BSW-O-Mat dataset.

Following the Michigan model (Campbell et al., 1960), core predictors of voting behaviour should remain stable across (short) time spans. However, in the context of self-selection biases, it may be that convergence validity cannot be achieved if samples differ too much. Although differences in relative group size are less significant, a VAA may reach only specific voters of a given party, particularly specialised VAAs. For example, if populist users do not use VAAs, maybe due to the lack of trust in political scientists, a voting effect related to populism cannot be detected with VAA datasets. Failing to address such self-selection biases can have drastic consequences, as the Literary Digest fiasco of 1936 illustrates (Squire, 1988).

Although studies have reported convergence validity between probability and VAA samples (Toshkov & Romeijn, 2021), the BSW-O-Mat VAA presents a more challenging test case. The tool was developed by political scientists from the University of Potsdam and is based on policies directly extracted from BSW's first manifesto. Single-party VAAs differ from classic VAAs in their selection mechanisms. While all VAAs overrepresent politically engaged users, classic multi-party VAAs attract users seeking general election guidance. Single-party VAAs pose a methodological challenge for testing convergent validity for three reasons. Firstly, they attract users specifically interested in that party, creating an additional self-selection bias compared to classic VAAs. Secondly, all policy items are drawn from the party's manifesto, which can affect the analysis of congruence, defined as the extent to which BSW voters align with the policy positions of the party. The analysis may potentially inflate congruence effects, as users agreed with policies directly

derived from BSW's manifesto; in contrast, researchers have typically used established policy items in surveys. Thirdly, as users know they are evaluating their alignment with a specific party, this could introduce additional biases (e.g., potential voters may agree with the statements because they like the party). These features make convergence with probability samples less likely but more meaningful if achieved. Finally, even if convergence is observed, both datasets could yield similar results because they are subject to similar biases. This could occur when comparing two VAA datasets. To rule this out, the reference dataset should be a probability sample. Although far from perfect, a probability sample is the "gold standard" in sampling (Pekari et al., 2022).

One sample that fulfils these criteria is the rolling cross-sectional sample from the GLES. The dataset includes more than 8,000 individuals who were interviewed around the 2025 Bundestag elections and is publicly available (GLES, 2025). As the study also features a supplementary wave, including, among other variables, items on populism, the dataset includes a large set of comparable items. The entire study was conducted between January 6 and March 31, 2025. Thus, there is a temporal limitation affecting the convergent validity test: the BSW-O-Mat data were collected in early 2024, while the GLES data come from the 2025 election period. The majority of the 50,000 users visited the BSW-O-Mat website in the first quarter of 2024. This means that most BSW-O-Mat users responded to the items around one year earlier than the respondents in the GLES 2025 study.

To what extent are the two datasets comparable? At both points in time, in January 2024 and January 2025, the BSW was polling closely to its actual Bundestag election result (4.98%), as shown in Figure 1. The party's approval peaked in late Summer 2024, shortly before the three state elections in Thuringia, Brandenburg, and Saxony, where the BSW did exceptionally well and entered two state governments (Patton, 2025). Although a dataset collected in autumn 2025 would be much closer to the Bundestag elections, it is more likely that the potential BSW electorate would differ more strongly, as by then the number of potential BSW voters had almost doubled.

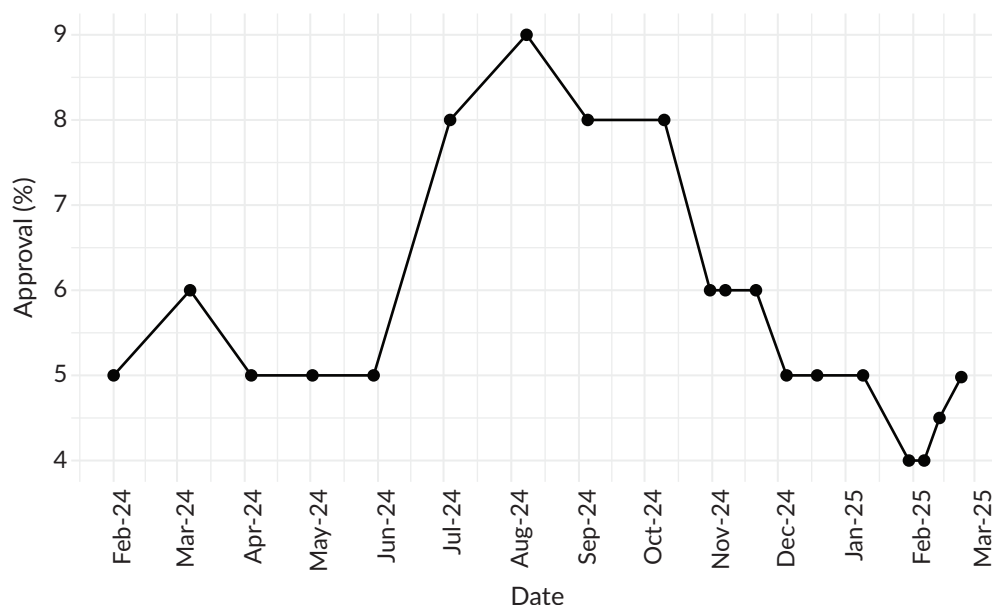


Figure 1. Approval of BSW in Infratest polls. Source: Zicht & Cantow (2025).

Another potential problem is that the electorate may have changed between 2024 and 2025, which may have been the case regarding age (see the previous section). Nevertheless, exit polls from the 2024 European and 2025 Bundestag elections reveal that BSW voter profiles remained stable: programmatic voting dominates, though dissatisfaction with other parties and the appeal of leader Wagenknecht matter substantially. Key policy priorities—foreign policy positions (ending weapon deliveries to Ukraine, “securing peace”), left-wing economics combined with immigration restriction—remained consistent across both elections (Infratest dimap, 2024, 2025a).

Both exit polls confirm that the party is mainly supported by those with medium levels of education, and to a lesser extent by those with high and low levels of education. BSW is also less popular among men than among women, likely because the party leader is female. The only observable shift reported is the reversal of the age effect, as discussed in the previous section. Overall, these exit poll results indicate that the profile of BSW voters was rather similar in early 2024 and early 2025.

5. Data Cleaning, Sampling, and Weighting

Several steps were taken to improve data quality through data cleaning. As a first step, all respondents with missing data for the independent and dependent variables of interest were removed. In the GLES dataset, the speeder index was used to remove speeders (i.e., the 10% fastest respondents, see codebook in GLES, 2025). In the BSW-O-Mat dataset, speeders and skimmers were identified based on the minimum time it took to read the items (Andreadis, 2014); respondents below that threshold were removed. In addition to the main VAA items, respondents could complete a background questionnaire on sociodemographics (before the main task started) and answer additional questions (e.g., attitudes such as populism, trust, and voting intention) after the match was calculated. Around 60% of the respondents completed the last question of the additional survey. The data were weighted by age, education, gender, and region. Those with missing data for these variables were removed, including those below the voting age (18), students at school (due to the lack of clear ISCED levels), and those not identifying as male or female (as the data are not present in the census). Respondents were grouped into four age categories: 18–34, 35–49, 50–65, and 65 and older. Population data were obtained from CensusHub (Eurostat, 2021). Concerning education, an indicator based on low education was built (0 = ISCED 1 or 2, i.e., no degree, a Hauptschule or Realschule degree, else = 1), as this allowed for the identification of respondents’ level of education in the absence of questions on further job qualifications. The region was coded by assigning states to East, West, or Berlin, as the city is composed of former Eastern and Western districts.

After data cleaning, 5,159 cases out of the initial 8,562 respondents remained for analysis in the GLES dataset. Although the dataset provides its own weights, the data were re-weighted using the exact same mechanism as for the BSW-O-Mat data (post-stratification weights based on known joint marginal distributions for the four variables). The full BSW-O-Mat dataset consists of 50,711 cases. After data cleaning, 7,537 respondents remained in the dataset. As around 93% of all BSW-O-Mat users visited the tool before the European elections (June 9, 2024), respondents after the EP elections were removed. As the number of respondents was still higher than in the GLES dataset, a quota sample with the same sample size as the GLES dataset (5,159) was drawn from the weighted dataset, where the chance of being selected was equivalent to a respondent’s weight. This ensured comparable statistical power across datasets, preventing the VAA’s larger sample from producing artificially higher significance rates that could affect the convergence validity test. Subsequently, the dataset

was reweighted to obtain the analysis weights. Weights ranged from 0.42 to 5.16 (BSW-O-Mat) and 0.08 to 5.0 (GLES), indicating moderate adjustment needs.

The main dependent variable, the intention to vote for BSW in the next Bundestag election, was coded as a binary indicator: 1 for BSW voters and 0 for all others; invalid voters and those who had not yet decided were removed from the analysis. Table 1 compares the composition of potential voters in both datasets, in terms of recall and voting intention, against the backdrop of the actual results. While it should be noted that the BSW-O-Mat data were collected around one year earlier, when Bündnis '90/Die Grünen were polling slightly better and Die Linke worse compared to their 2025 election result, it becomes clear that the BSW-O-Mat dataset is strongly skewed towards potential BSW voters, which raises doubts about convergence validity with the probability sample results. A full 34.3% of users indicated that they would vote for the new party, compared with only 6.0% of GLES respondents. AfD, SPD, and CDU/CSU voters are strongly underrepresented in the VAA dataset, as are non-voters; those indicating an intention to vote for Die Linke are overrepresented. Nevertheless, neither does the voting intention in GLES accurately reflect the election results—even here, BSW voters are overrepresented, just as potential voters of the AfD, CDU/CSU, and SPD. GLES also fails to reach a substantial share of non-voters. These trends are also visible for the 2021 recall questions, although the BSW-O-Mat data are even more strongly skewed towards former Bündnis '90/Die Grünen and Die Linke voters.

Table 1. Voting intention and vote recall in the two datasets (unweighted distributions).

Party	Voting Intention 2025		Result 2025	Recall 2021		Result 2021
	GLES (2025)	BSW-O-Mat (2024)		GLES (2025)	BSW-O-Mat (2024)	
BSW	6.0	34.3	4.1			
AfD	19.8	5.5	17.2	12.5	4.4	8.0
CDU/CSU	29.2	13.9	23.5	25.1	10.2	18.5
FDP	4.1	5.0	3.6	8.0	11.9	8.8
Bündnis '90/ Die Grünen	11.0	18.9	9.6	12.8	25.0	11.4
Die Linke	5.3	10.2	7.2	3.9	21.8	3.8
SPD	19.2	7.0	13.5	26.5	16.2	19.8
Other	2.9	4.0	3.7	2.8	5.1	6.0
No right to vote				2.5	2.5	
Non-voters	2.4	1.1	17.6	6.0	2.7	23.7
<i>n</i>	5,159	5,159		5,159	5,159	
		(quota sample)			(quota sample)	

Note: As only around 74% of the German population were eligible voters, but both datasets had eligible voters as their target population, the 2021 election results are reported as percentages of eligible voters.

While the fact that VAA data, especially data from such specialised VAAs, are skewed is not surprising, the summary in Table 1 also indicates the bias's direction: voters from the left side of the political spectrum are overrepresented in the BSW-O-Mat dataset. Nevertheless, this does not automatically risk convergence validity. The self-selection bias in VAA data is a function of multiple factors, and when a VAA offers a match with only one specific party, some voter groups are more likely to be interested in this tool than others.

The key question is whether biases also extend beyond the relative size of groups, i.e., systematic biases linked to voters' preferences. Regarding the BSW, this could be the case if the BSW-O-Mat only reaches, for example, liberal BSW voters, which would reduce observed convergence.

A more nuanced picture of BSW-O-Mat users emerges when examining the distribution of the propensity-to-vote (PTV) variable for BSW, which was included in the BSW-O-Mat tool but not in GLES. Typically, PTV variables are skewed towards negative evaluations, even for mainstream/catch-all parties (A. Wagner, 2017). Figure 2 shows the PTV distribution for the BSW-O-Mat users. The best and worst evaluation categories stand out as those that were used most frequently. In line with the voting intention reported in Table 1, the website attracted more users with a strong propensity to vote for BSW, but those evaluating BSW negatively were also present, thereby increasing polarisation. This could lead to larger effect sizes, as there is a tendency to compare unlikely BSW voters to likely BSW voters.

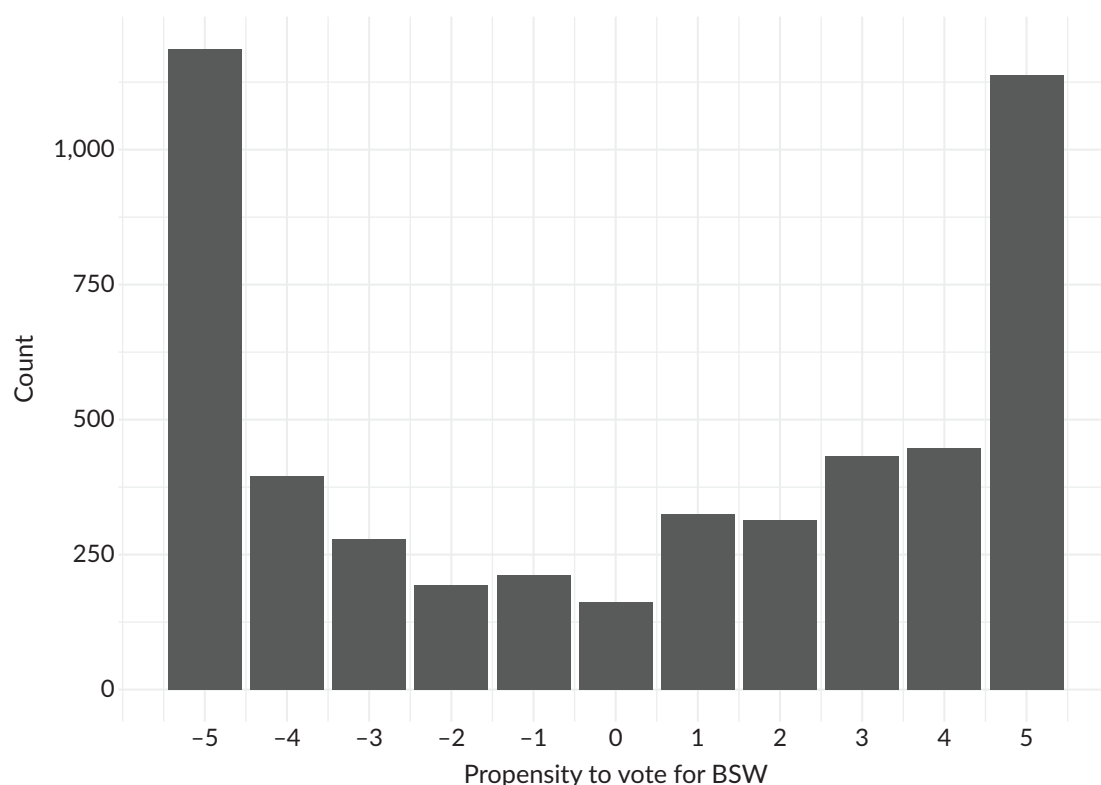


Figure 2. Univariate distribution of propensity to vote for BSW in the BSW-O-Mat dataset (unweighted).

5.1. Independent Variables

Based on the Section 3 literature review, independent variables are organised into four groups: socio-demographic factors, policy preferences, trust and populism, and past voting behaviour. The relatively large set of comparable items was a key criterion for selecting GLES as the comparison dataset. The independent variables of interest can be summarised into four groups: socio-demographic factors, policy preferences, trust and populism, and past voting behaviour. Age (18–34, 35–49, 50–65, and 65+), gender (male/female), education, and region (East/West/Berlin) are the socio-demographic factors that will be tested. Whilst the primary goal is testing convergence rather than causal hypotheses, the Section 3 literature

suggests expected patterns: East Germans and women are more likely to support BSW, and university graduates are less likely. The age effect is likely to differ across the two datasets, as discussed in Section 3.

Regarding policy preferences, the BSW-O-Mat dataset offers a much richer set of variables, as all 27 main VAA items are policies proposed by the BSW in its manifesto. BSW-O-Mat items measure user policy positions that, because they derive directly from BSW's manifesto, simultaneously operationalise congruence with BSW. In contrast, GLES items measure general policy positions, requiring us to calculate congruence separately by comparing responses to BSW's party positions. This could hamper the identification of similar items in the GLES dataset. Six items in the GLES dataset were identified as measuring similar constructs: climate change, weapon deliveries to Ukraine, taxation, economic intervention, immigration, and integration (see Table A1 in the Supplementary File). To additionally analyse congruence with the GLES dataset, congruence was calculated for four of these items (climate change, economic intervention, immigration, and integration) using BSW party positions from the GEPARTEE expert survey on political parties in Germany (Thomeczek et al., 2024). Here, congruence is defined as the absolute distance from BSW's position by the GEPARTEE, based on a simple proximity model. As the GLES item on weapons for Ukraine reflects BSW's direct position (Bündnis Sahra Wagenknecht, 2025, pp. 6–7), agreement with this item was regarded as congruence with BSW's position. Thus, policy congruence with BSW's positions was calculated for a total of five items in the GLES dataset, which was tested in an additional model.

To measure populist attitudes, the BSW-O-Mat dataset included items from the populism scale developed by Castanho Silva et al. (2019), while the GLES dataset included six items from various populism scales. Six items from the BSW-O-Mat dataset were selected, as they share the common “people vs. elite” denominator with the GLES items. In addition, trust in the federal government, measured using an 11-point Likert scale, was included in both datasets. Satisfaction with democracy in Germany was measured using a 5-point Likert scale in both datasets. The final set of items addresses vote switchers. Based on the recall question for the 2021 Bundestag elections, the extent to which the different voter groups support the BSW will be analysed. Based on the literature review, the focus lies on Die Linke, AfD, and non-voters.

6. Analysis

6.1. Predicting BSW Vote Intention

The analysis addresses convergence validity, specifically whether models based on the two datasets yield similar conclusions. For substantial effect convergence to be present, statistically significant effects should point in the same direction, while the second part of the analysis focuses on effect size convergence. Due to the nature of the VAA, effect sizes may differ from those in the GLES probability sample. Generally, it is expected that the effect of the policy items will be stronger in the BSW-O-Mat analysis, as these items are derived directly from the BSW manifesto, which was also transparent to users. Nevertheless, apart from the age effect, which appears to have been reversed between early 2024 and early 2025, this analysis remains based on the assumption that the core BSW electorate regarding the tested factors is relatively stable and has not undergone substantial change.

Prior to the analysis, all variables were recoded to have the same direction. The policy items were recoded so that agreement indicated support for policies associated with BSW's ideology. Numeric variables were

transformed to a range of 0 to 1 (min-max scaling) to facilitate comparison with variables of different scales. The results are presented as average marginal effects (95% CIs) from the weighted logistic regression models. Average marginal effects have the advantage over logistic regression coefficients in that they allow for direct comparisons between variables and, in this case, across datasets. They indicate the average marginal change in the probability of voting for the BSW across all cases, where the marginal effects differ for respondents, unlike in linear models.

The first model, which includes the socio-demographic factors, is presented in Figure 3. The models from both datasets show a higher average marginal probability of voting for BSW among East Germans and female voters, and a lower probability among those holding a university degree. Regarding age, the effect points in opposite directions, as expected. Generally, the average marginal effects are stronger in the model based on the BSW-O-Mat dataset. The socio-demographic variables are kept as additional variables in the following models to test their robustness.

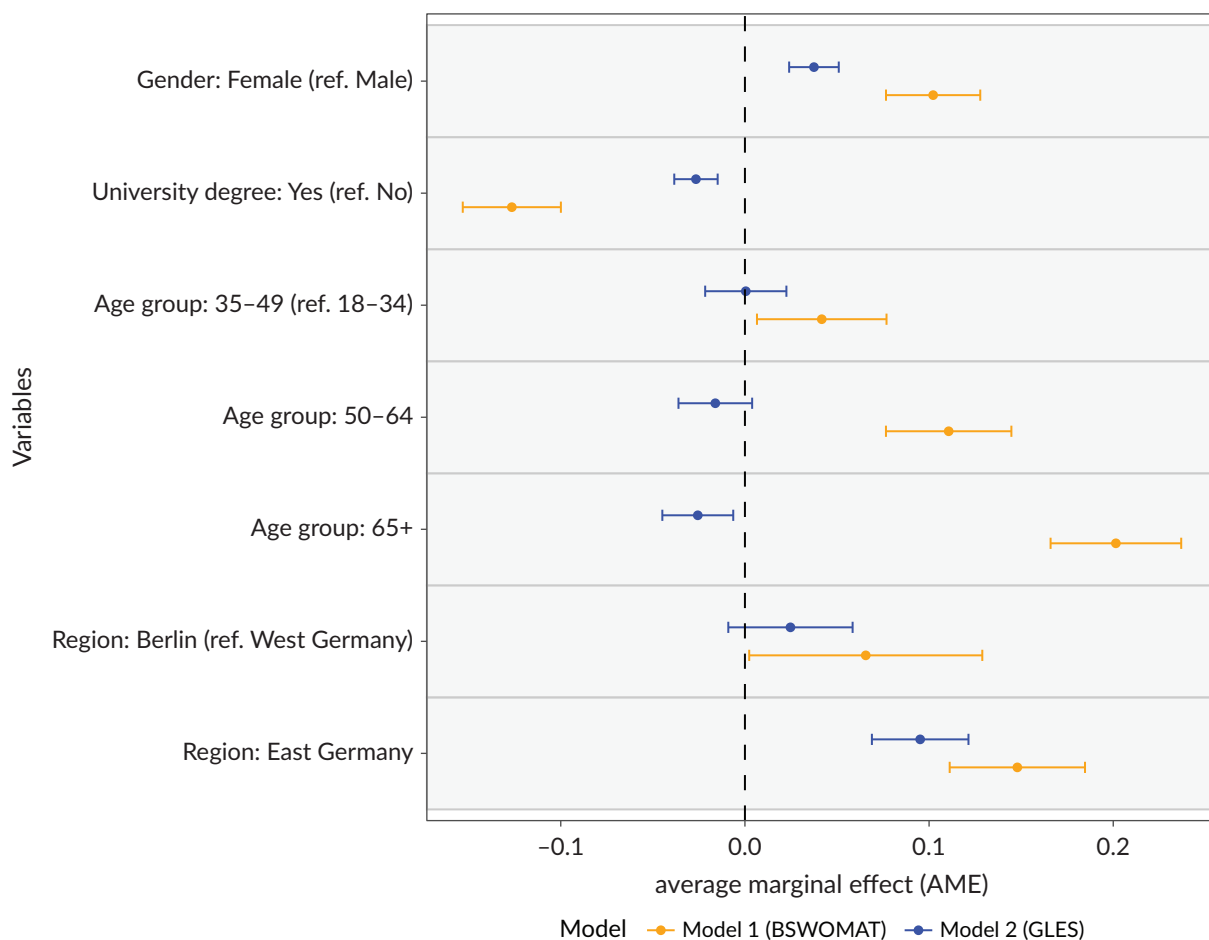


Figure 3. Average marginal effects (95% CIs) of weighted regression models predicting a BSW vote for socio-demographic factors.

The effects of the policy-related factors are summarised in Figure 4. Two models were estimated using the GLES dataset: one for policy support (blue) and one for policy congruence based on the BSW position from the GEPARTEE project (red). The BSW-O-Mat models, which measure support for BSW policies using BSW's

wording, report strong and statistically significant effects for all policy-related items. Unsurprisingly, the effect is weakest for the item that does not originate from the main BSW-O-Mat VAA items (i.e., a policy proposal not derived from the BSW manifesto), namely the item on state intervention in the economy, which was asked in the additional survey. Turning to the GLES analysis, the effects of supporting higher taxation of the wealthy, favouring state intervention in the economy, and rejecting weapon deliveries to Ukraine all point in the same direction. The GLES dataset contains two items related to immigration: immigrants' adaptation to German culture (assimilation) and the limitation of immigration. Surprisingly, the assimilation item shows no effect, whilst the immigration limitation item shows a negative effect. This likely reflects measurement issues: the assimilation item shows near-ceiling agreement across all voters (the mean is greater than 0.6 even among the least supportive group, which is voters of the Bündnis '90/Die Grünen; see Table A3 in the Supplementary File), resulting in insufficient variance. The immigration limitation item may capture general restrictionism rather than BSW's specific position, which emphasises cultural integration alongside such limitations. The congruence model addresses this by measuring proximity to BSW's exact stance, yielding a significant positive effect. No significant effect is reported for the climate-versus-economy item.

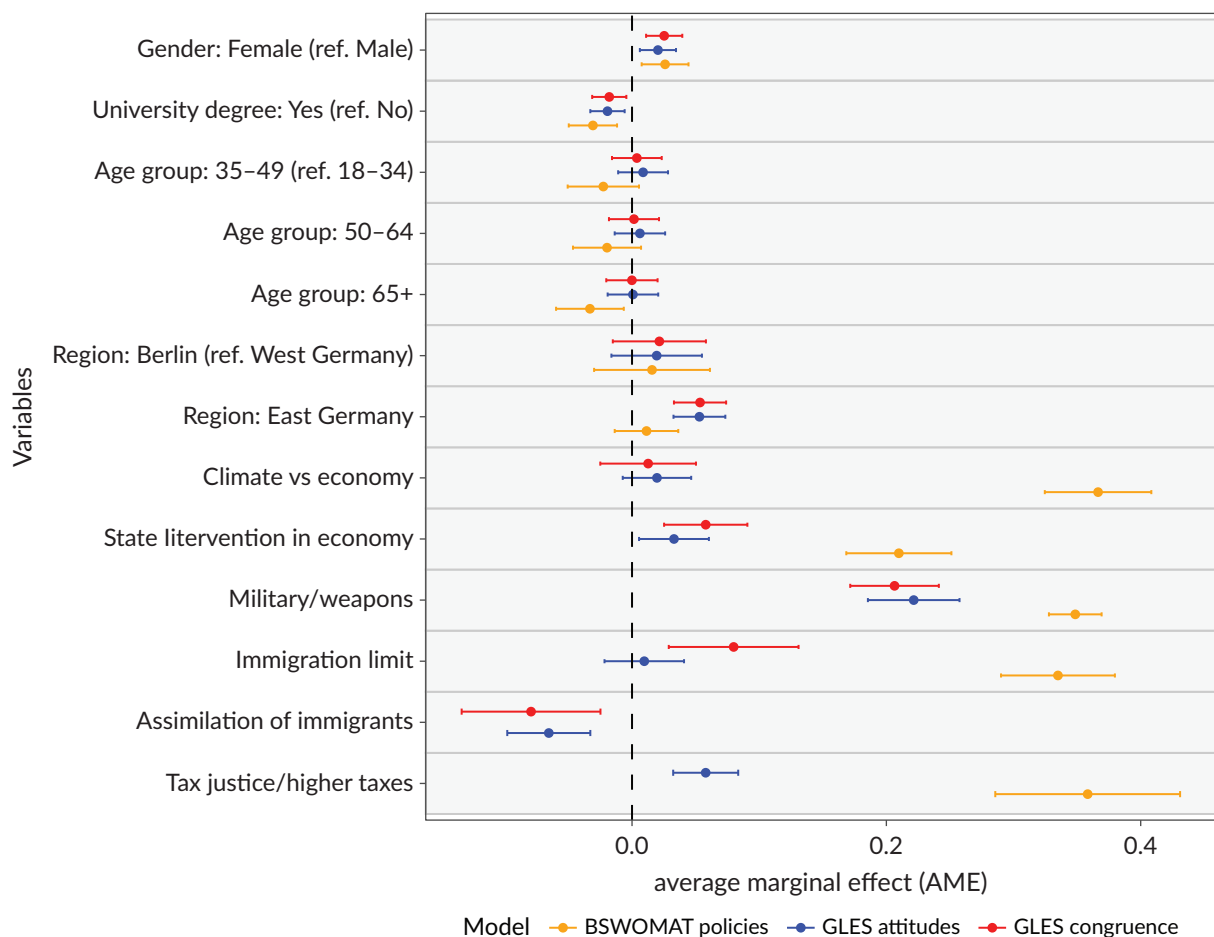


Figure 4. Average marginal effects (95% CIs) of weighted regression models predicting a BSW vote for policy-related factors.

For four variables (state intervention, climate versus economy, immigration limit, and weapon supply to Ukraine), congruence could be calculated for the GLES dataset, as reported in the second model (red). Overall, the effects are stronger compared to pure policy support, and the immigration item also shows a

statistically significant effect. Regarding immigration, it seems logical that congruence matters more than raw policy support, as the BSW advocates for limiting immigration but takes a less radical stance compared to the AfD.

The results for populism and institutional trust are summarised in Figure 5. As discussed in Section 5.1, the populism items differed, which prevented the formation of a comparable index. However, multicollinearity does not pose a problem in the two models—the VIF values for all populism variables are smaller than 2. Both datasets reveal consistent patterns. Although the set of populism items varied, the results from both models indicate that agreeing with these items generally increases the likelihood of voting for the BSW. In contrast, trusting the federal government and being satisfied with contemporary democracy in Germany lowers this likelihood. Again, a trend is visible that the effects are larger in the analysis with the BSW-O-Mat dataset.

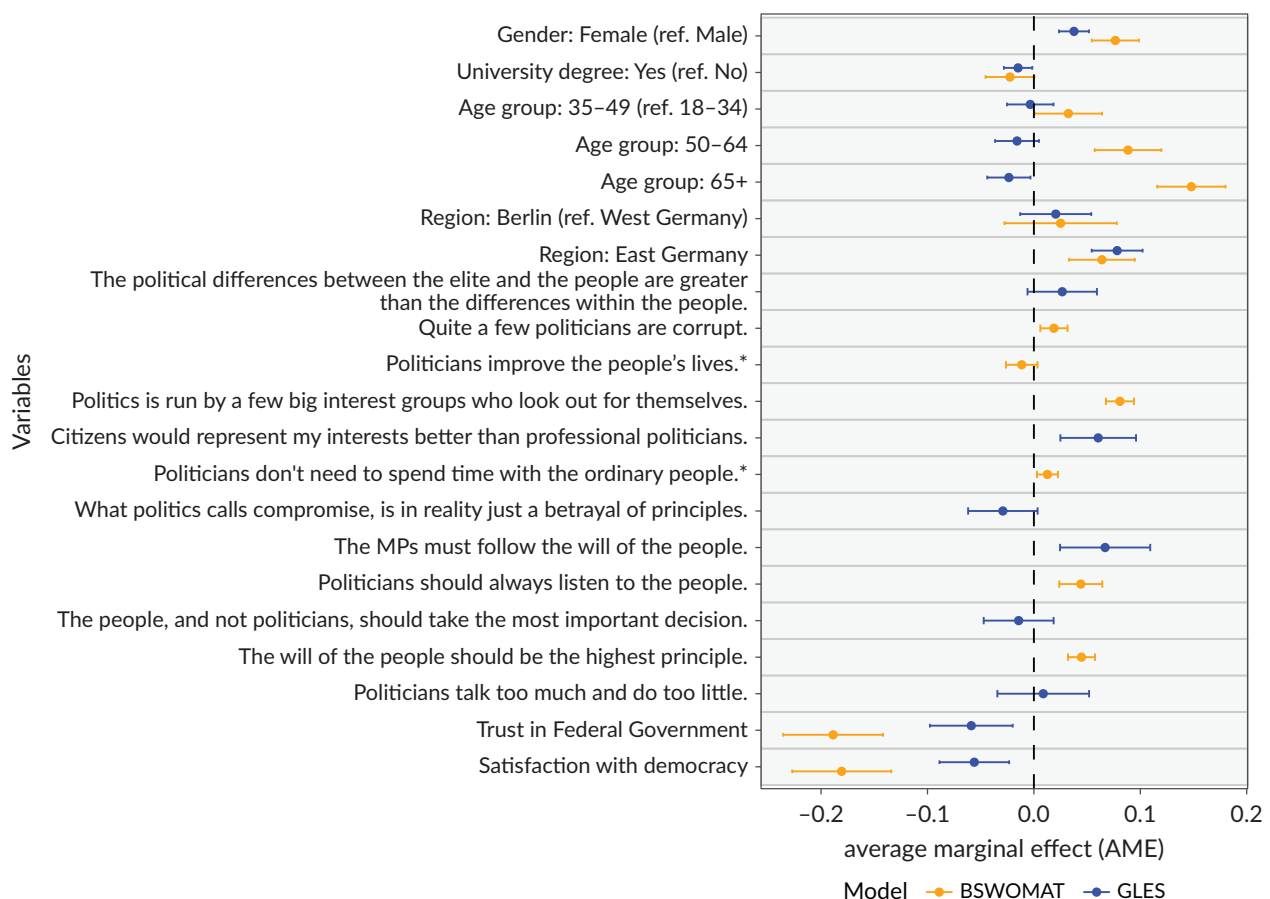


Figure 5. Average marginal effects (95% CIs) of weighted regression models predicting a BSW vote for institutional trust and populism.

Finally, Figure 6 shows the average marginal effects for voting BSW across three voter groups with high expected BSW potential: former AfD voters, Die Linke voters, and non-voters. Both models show a significant positive effect for former non-voters and Die Linke voters. The effect for non-voters is much stronger in the BSW-O-Mat model. This is likely due to additional selection bias: non-voters using the website had greater affinity for BSW than non-voters in the representative sample. Regarding AfD voters from 2021, the BSW-O-Mat model reports no significant effect. The average marginal effect is almost the

same magnitude in the GLES model but is estimated with less uncertainty, resulting in a significant, albeit minimal, negative effect (the upper boundary of the confidence interval is -0.001). Therefore, it can be concluded that the BSW was able to draw strongly from former Left voters, some non-voters, and, contrary to initial expectations (S. Wagner et al., 2023), only minimally from former AfD voters, if at all.

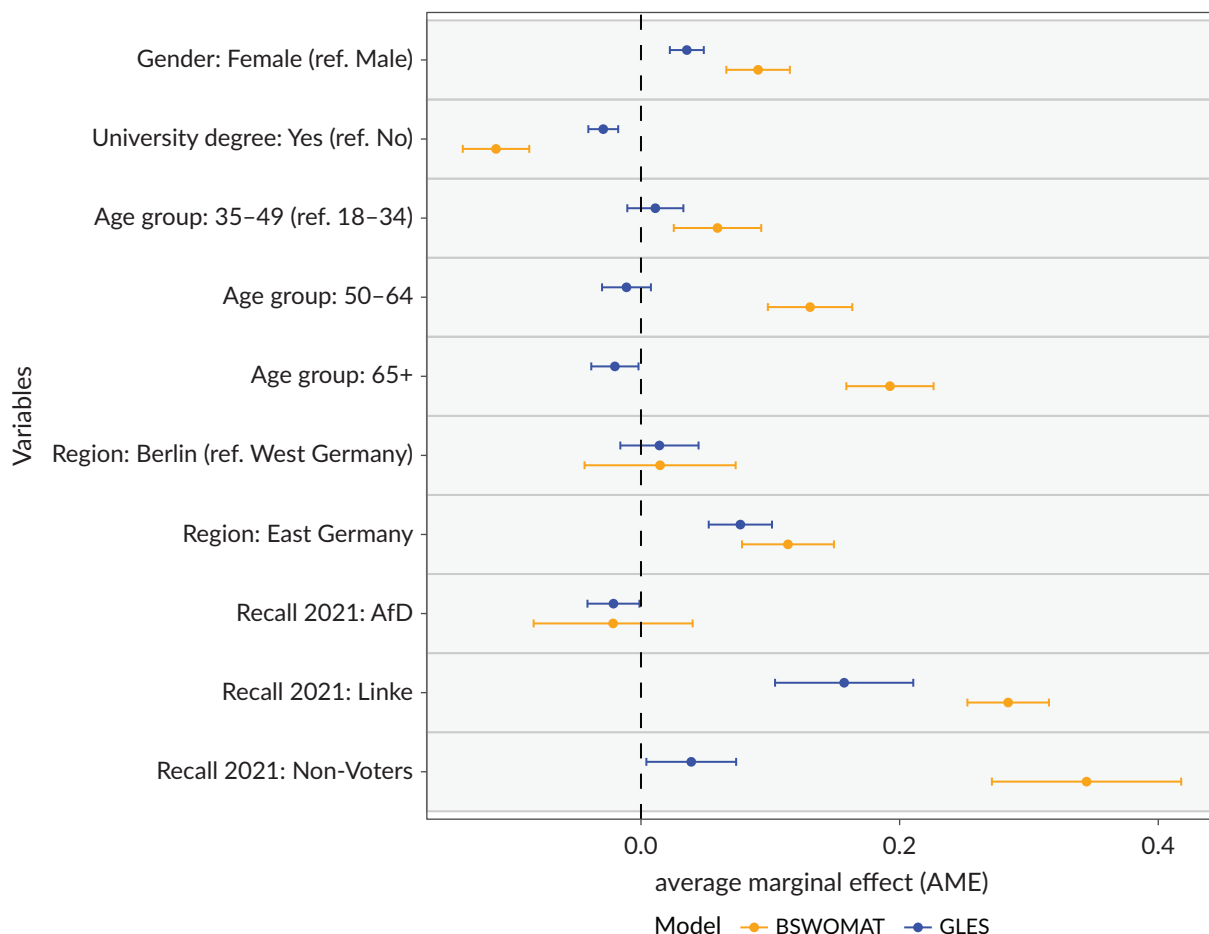


Figure 6. Average marginal effects (95% CIs) of weighted regression models predicting a BSW vote for recall.

To further analyse effect size convergence, Figure 7 illustrates the correlation between the estimated average marginal effects in both models, excluding the populism variables (as the items differed) and the effects for age (since the effects, as expected, pointed in different directions). Among all effects, the correlation is moderately high with $r = 0.654$. In addition, all estimates are either positive in both models or negative in both models, i.e., they do not point in different directions. As indicated by the colour, some outliers concern effects that were only statistically significant in one of the models. The right plot shows that removing these pairs increases the correlation to $r = 0.872$, indicating a strong correlation.

Lastly, because the largest weight in the BSW-O-Mat dataset was approximately 12 times that of the smallest, the average marginal effects were recalculated using unweighted data to assess the robustness of the results relative to the original 8,562 respondents. The estimated effects only deviate minimally. The correlation between the average marginal effects from the unweighted analysis and the weighted BSW-O-Mat analysis is very high, with $r = 0.991$ (see Figure A1 in the Supplementary File). Thus, while

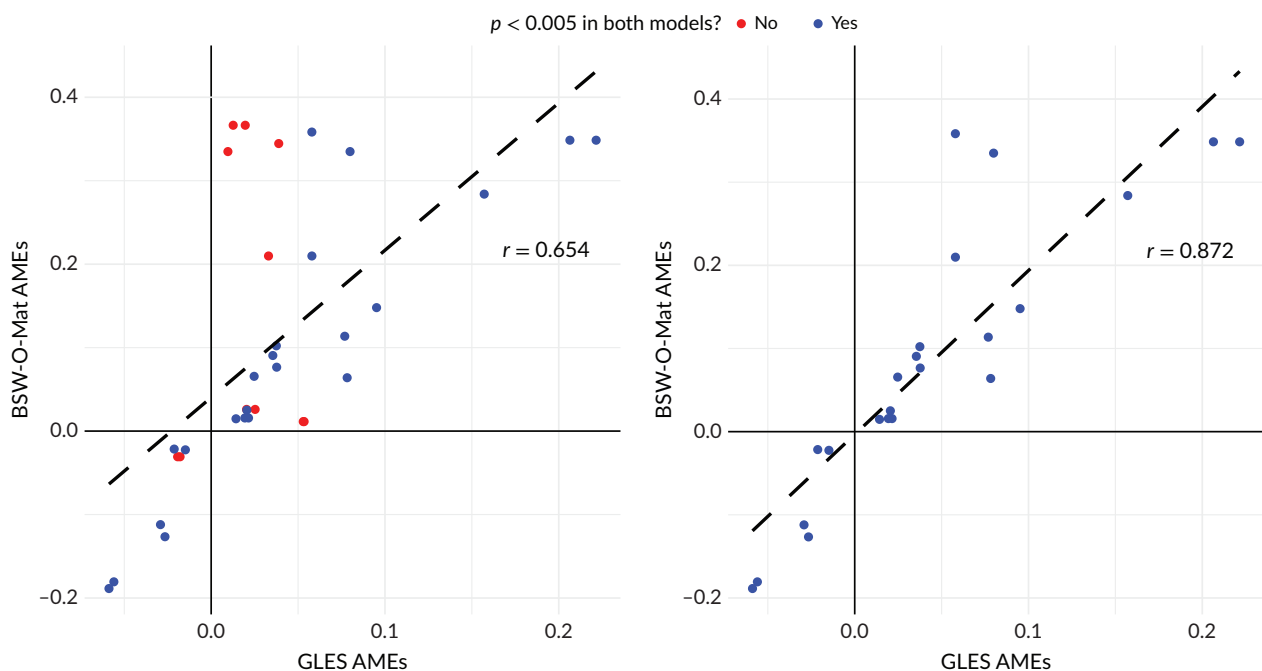


Figure 7. Correlation of average marginal effects in GLES and BSW-O-Mat models.

BSW-O-Mat users are demographically not representative, their response patterns within demographic groups remain stable after weighting.

6.2. Robustness Test: Predicting Underrepresented AfD Voters

Section 6.1 has demonstrated that the BSW-O-Mat dataset is, overall, useful for analysing BSW support. In this section, the analysis will be repeated for predicting AfD support. As shown in Table 1, AfD voters are strongly underrepresented in the BSW-O-Mat dataset. While this alone does not pose a problem—it could be a bias only affecting the relative size of voter groups—the fact that the VAA was designed to analyse congruence with the BSW could have attracted only a specific segment of AfD voters, namely those interested in the BSW, which could be another self-selection bias. To test whether the BSW-O-Mat dataset can also predict voting for the AfD, the same models were rerun for both datasets, with AfD voting intention as the dependent variable. Regarding recall, the largest voter groups switching to the AfD were former CDU/CSU voters, FDP voters, and non-voters (Infratest dimap, 2025b). As the AfD (unlike the BSW) also competed in the 2021 elections, the model includes a dummy variable for 2021 AfD voters.

Despite AfD voters constituting only 5.5% of the BSW-O-Mat sample (vs. 19.8% in GLES), substantial effect convergence is achieved (see Figures A2–A5 in the Supplementary File), suggesting that the tool captures general political dimensions rather than BSW-specific artefacts. The successful AfD prediction strengthens confidence in the BSW results, as it demonstrates that the dataset's validity extends beyond its intended focal party.

Both datasets show that the AfD is less supported by women, West Germans, and those with a university degree, and is more supported by the 35–49 age group than by the youngest age group. The policy models show that AfD voters prefer less state intervention in the economy, less taxes, the assimilation of migrants,

and a strict immigration limit, and are critical of weapon deliveries to Ukraine. The last two effects (immigration limitation and weapon deliveries) were not significant in the BSW-O-Mat analysis. This is expected: BSW-O-Mat items measure congruence with BSW's positions (opposing all military intervention; moderate immigration restriction), not AfD's positions (no strict anti-militarism; radical immigration restriction). The GLES items capture more general stances, aligning more closely with AfD's policy demands (Alternative für Deutschland, 2025), thus showing significant effects. The populism and trust models reveal a significant negative relationship between trust in government and satisfaction with democracy. Regarding populism, surprisingly, there are only a few significant positive effects across models in items related to the people's will, and even statistically significant negative effects. It should be noted that the GLES populism battery includes several new items that have not been validated yet. It was confirmed that FDP and CDU/CSU voters in 2021 were more likely to vote for the AfD. For non-voters, both the GLES and BSW-O-Mat models show a positive effect; however, the lower boundary is -0.0015 in the latter, indicating an effect that is not significantly different from 0. This is another indication that non-voters closer to BSW policies are overrepresented.

7. Conclusions

As the VAA market has become increasingly crowded, specialised VAAs focused on specific topics or parties have become more popular recently. The usefulness of such VAA data has not yet been studied; however, they may be subject to additional biases related to the tool's topic. Therefore, such VAA datasets present an even harder case for testing convergence validity with probability samples. This article has focused on the specialised single-party VAA BSW-O-Mat, developed in early 2024 to evaluate a match with the newly founded party BSW. The VAA dataset was then compared with the 2025 election probability sample from the GLES. Despite a strong overrepresentation of potential BSW voters, this study demonstrates that single-party VAA data can achieve convergent validity with probability samples. Convergence validity of substantial effects was observed in almost every case. The two main differences between the BSW-O-Mat and the GLES models were related to effect sizes, which were generally stronger in the former, and the robustness of some socio-demographic factors.

Convergence validity was put to an even harder test by running the same models used to predict voting intention for the AfD, whose voters were strongly underrepresented in the BSW-O-Mat dataset. The results for AfD voters, surprisingly, showed a similarly high level of convergence. Overall, these results contribute substantial insights into current methodological debates in VAA research. They suggest that such single-party VAAs can yield valuable (and valid) insights into the study of new parties, offering a promising avenue for researchers focusing on new parties or specific topics. This is generally good news for VAA researchers and developers, as the data have proven to be useful even for strongly underrepresented voter groups.

The (strong) self-selection bias—a challenge for almost every VAA dataset—present among potential BSW voters did not compromise convergence validity. The nature of the VAA (a tool focusing on policies proposed by a single party) is likely the reason for differences in effect sizes: First, because it attracted more BSW voters; second, because all statements were derived directly from the party's manifesto. The state intervention item demonstrates that this is the only policy item asked identically in both datasets and shows the strongest effect-size convergence among policy variables. Nevertheless, larger effect sizes were also observed for other variables, such as trust, education, and gender. For VAA researchers, this means being

cautious when interpreting effect sizes, but they can be relatively confident in the direction and statistical significance of effects.

These findings advance VAA methodology in three ways. Firstly, they establish that convergence validity—previously demonstrated only for classic multi-party VAAs (Toshkov & Romeijn, 2021)—extends to single-party tools despite compounded selection bias. Secondly, they clarify that substantial-effect convergence is more robust than effect-size convergence, guiding researchers as to when VAA data support explanatory inferences and when caution is warranted in interpreting magnitudes. Thirdly, the successful AfD analysis demonstrates that VAA data validity extends beyond focal voter groups, enabling comparative analyses within single-party VAA datasets.

Several limitations emerged from the analysis, which have implications for VAA research. Firstly, VAA designers need to know and address these biases (Pianzola, 2014). Effects related to non-voters differed most strongly and, in the case of the AfD models, could not be replicated. While the tool attracted a substantial share of non-voters, it is generally not suited to study this group, as it apparently attracted predominantly non-voters interested in the BSW. Secondly, convergence validity may have been facilitated by BSW's large public attention, which may have attracted broader segments. Other tools for new parties may attract fewer, more self-selected users. Thirdly, because VAA items were directly derived from the party's manifesto, it was more difficult to compare the results on issue congruence with those from the GLES study. In the future, VAA designers could consider using established survey items in these contexts as well. However, this may pose a challenge for new parties, as it is often unclear to researchers where such parties stand. Finally, there is the obvious limitation that the data covered only one specialised VAA in Germany, but it would be interesting to compare it with similar new-/single-party VAAs.

The analysis has revealed the potential of data from specialised VAAs for future research. While such tools have been around for some time, their data are hardly analysed. The fact that a specialised VAA for a new party, such as BSW, attracted so many interested users and yielded analytical insights comparable to those from probability samples is reassuring for VAA developers. Most countries have seen new parties emerge that quickly gained support (e.g., Dutch New Social Contract; Otjes & De Jonge, 2024). Having similar tools for more new parties can not only help voters understand what these new parties stand for, but also provide initial insights for political science on the electoral potential of such parties without requiring the wait or investment in probability samples. This also reveals the democratic function of such tools, as they can help voters to inform themselves about new parties or specific topics.

Another possibility is to view specialised VAAs as tools for analysing hard-to-reach groups. While such scenarios are thinkable (e.g., migrants via multilingual VAAs), the BSW analysis does not test this application, as BSW voters are not systematically underrepresented in probability samples. While this study compared single-party VAA data to probability samples—the methodological gold standard—future research could complement this by comparing single-party VAAs to classic multi-party VAAs to isolate VAA-specific versus single-party-specific selection mechanisms. Such comparisons would clarify which biases stem from VAA usage in general versus specialised tool design.

Acknowledgments

I wish to thank my colleague Aiko Wagner for his valuable feedback.

Funding

The hosting of the website BSW-O-Mat was kindly supported by the alumni club of the University of Potsdam (Universitätsgesellschaft Potsdam e.V.).

Conflict of Interests

The author declares no conflict of interests.

Data Availability

After the end of the projects associated with the dataset, the data will be made publicly available at <https://doi.org/10.7910/DVN/PGIGYC>

LLMs Disclosure

For support with R code, overall feedback, and language check, Perplexity.AI was used.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529–546. <https://doi.org/10.1017/S0003055401003100>
- Alternative für Deutschland. (2025). *Zeit für Deutschland*. https://www.afd.de/wp-content/uploads/2025/02/AfD_Bundestagswahlprogramm2025_web.pdf
- Andreadis, I. (2014). Data quality and data cleaning. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting advice applications in a comparative perspective* (pp. 79–92). ECPR Press.
- Baumann, H., & Seikel, D. (2024). *Wer wählt "Bündnis Sahra Wagenknecht"? Aktuelle Erkenntnisse über die BSW-Wahlbereitschaft aus der WSI-Erwerbspersonenbefragung* (Research report No. 94). WSI Report. <https://www.econstor.eu/handle/10419/299270>
- Bundesarchiv. (2018). *Wahl zur Nationalversammlung—1919: Wie hätten Sie gewählt?* <https://weimar.bundesarchiv.de/static/wahlkompass/index.html>
- Bundeswahlleiterin. (2024). *Europawahl 2024. Wahlbeteiligung und Stimmabgabe nach Geschlecht und Altersgruppen*. Bundeswahlleiterin/Statistisches Bundesamt.
- Bundeswahlleiterin. (2025). *Bundestagswahl 2025. Wahlbeteiligung und Stimmabgabe nach Geschlecht und Altersgruppen*. Bundeswahlleiterin/Statistisches Bundesamt.
- Bündnis Sahra Wagenknecht. (2025). *Unser Land verdient mehr! Wahlprogramm für die Bundestagswahl 2025*. <https://bsw-vg.de/wp-content/themes/bsw/assets/downloads/BSW%20Wahlprogramm%202025.pdf>
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. John Wiley.
- Castanho Silva, B., Andreadis, I., Anduiza, E., Blanusa, N., Morlet Corti, Y., Delfino, G., Rico, G., Ruth-Lovell, S., Spruyt, B., Steenbergen, M., & Littvay, L. (2019). Public opinion surveys: A new scale. In K. A. Hawkins, R. Carlin, L. Littvay, & C. Rovira Kaltwasser (Eds.), *The ideational approach to populism: Theory, method & analysis* (pp. 150–178). Routledge
- Eurostat. (2021). *The 2021 population and housing censuses in the EU*. <https://ec.europa.eu/eurostat/documents/4031688/14081269/KS-09-21-344-EN-N.pdf/5907978a-011d-52fc-100e-f6a67735d938?t=1641392358489>

- Ferreira da Silva, F., Reiljan, A., Cicchi, L., Trechsel, A. H., & Garzia, D. (2021). Three sides of the same coin? Comparing party positions in VAAs, expert surveys and manifesto data. *Journal of European Public Policy*, 1, 150–173. <https://doi.org/10.1080/13501763.2021.1981982>
- Garzia, D., & Marschall, S. (2019). Voting advice applications. In W. R. Thompson (Ed.), *Oxford research encyclopedia of politics*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228637.013.620>
- 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>
- German Longitudinal Election Study. (2025). ZA10101 GLES Rolling Cross-Section 2025ZA10101 GLES Rolling Cross-Section 2025 (Version 2.0.0) [Dataset]. GESIS Data Archive. <https://doi.org/10.4232/5.ZA10101.2.0.0>
- Heckmann, L., Wurthmann, L. C., & Wagner, S. (2025). Who's afraid of Sahra—Understanding the shift in votes towards Germany's Bündnis Sahra Wagenknecht. *Research & Politics*, 12(1), Article 20531680241311504. <https://doi.org/10.1177/20531680241311504>
- Herold, M., & Otteni, C. (2025). Closing the left-conservative gap of political representation? Ideological positions and democratic attitudes of “Bündnis Sahra Wagenknecht” (BSW) supporters in Germany. *German Politics*. Advance online publication. <https://doi.org/10.1080/09644008.2025.2497083>
- Hoffmann, P. (2025). Between left-authoritarians, sympathy for Russia, and political discontent: The correlates of voting for the Bündnis Sahra Wagenknecht (BSW). *Politische Vierteljahresschrift*, 66, 805–828. <https://doi.org/10.1007/s11615-025-00600-2>
- Hooghe, M., & Teepe, W. (2007). Party profiles on the web: An analysis of the logfiles of non-partisan interactive political internet sites in the 2003 and 2004 election campaigns in Belgium. *New Media & Society*, 9(6), 965–985. <https://doi.org/10.1177/1461444807082726>
- Infratest dimap. (2024). *Europawahl 2024: Wer wählte das BSW—und warum?* tagesschau.de. <https://www.tagesschau.de/wahl/archiv/2024-06-09-EP-DE/umfrage-bsw.shtml>
- Infratest dimap. (2025a). *Bundestagswahl 2025: Wer wählte das BSW—und warum?* tagesschau.de. <https://www.tagesschau.de/wahl/archiv/2025-02-23-BT-DE/umfrage-bsw.shtml>
- Infratest dimap. (2025b). *Bundestagswahl 2025: Wie die Wähler wanderten.* tagesschau.de. <https://www.tagesschau.de/wahl/archiv/2025-02-23-BT-DE/analyse-wanderung.shtml>
- Jankowski, M. (2024). Das Schließen der Repräsentationslücke? Die Wählerschaft des Bündnis Sahra Wagenknecht—Eine Analyse basierend auf Paneldaten. *Politische Vierteljahresschrift*, 66, 357–379, <https://doi.org/10.1007/s11615-024-00583-6>
- Ladner, A., & Fivaz, J. (2012). Voting advice applications. In N. Kersting (Ed.), *Electronic democracy* (pp. 177–198). Verlag Barbara Budrich.
- Marschall, S. (2014). Profiling users. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates: Voting advice applications in comparative perspective* (pp. 93–104). ECPR Press.
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706. <https://doi.org/10.1080/10584609.2020.1843572>
- Otjes, S., & De Jonge, L. (2024). The Netherlands: Political developments and data in 2023. *European Journal of Political Research Political Data Yearbook*, 63(1), 317–336. <https://doi.org/10.1111/2047-8852.12436>
- Patton, D. F. (2025). The Sahra Wagenknecht Alliance (BSW): A new party launches in Germany. *Journal of Contemporary Central and Eastern Europe*, 33(3), 645–662. <https://doi.org/10.1080/25739638.2025.2534407>
- Pekari, N., Lipps, O., Roberts, C., & Lutz, G. (2022). Conditional distributions of frame variables and voting

- behaviour in probability-based surveys and opt-in panels. *Swiss Political Science Review*, 28(4), 696–711. <https://doi.org/10.1111/spsr.12539>
- 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>
- Sheppard, J. (2018). *Measuring social attitudes with voter advice application data* [CSRM & SRC Methods Paper No. 3/2018]. <http://hdl.handle.net/1885/261446>
- Squire, P. (1988). Why the 1936 literary digest poll failed. *Public Opinion Quarterly*, 52(1), 125–133. <https://doi.org/10.1086/269085>
- Steiner, N. D., & Hillen, S. (2021). Vote choices of left-authoritarians: Misperceived congruence and issue salience. *Electoral Studies*, 70, Article 102280. <https://doi.org/10.1016/j.electstud.2021.102280>
- Steiner, N. D., & Hillen, S. (2025). Voting for the Bündnis Sahra Wagenknecht (BSW) from a Policy Space Perspective. *German Politics*. Advance online publication. <https://doi.org/10.1080/09644008.2025.2502100>
- Thomeczek, J. P. (2024a). Is Bündnis Sahra Wagenknecht a new alternative for populist voters? *Populism*. Advance online publication. <https://doi.org/10.1163/25888072-bja10069>
- Thomeczek, J. P. (2024b). The voting potential of Bündnis Sahra Wagenknecht (BSW). *Political Studies Review*, 23(1), 409–417. <https://doi.org/10.1177/14789299241264975>
- Thomeczek, J. P., Wurthmann, C., & Stecker, C. (2024). *Europartycheck: VAA User Data and GEPARTEE Expert Survey (Party Positions)* [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/7PBJS9>
- Thomeczek, J. P., Zhirnov, A., Scotto di Vettimo, M., Buma, R., Cristea, V., Dolle, M., Kersting, N., Krouwel, A., Lange, S., Lopez Ortega, A., Thompson, J., Antonucci, L., Di Stefano, R., & Ruggeri, F. (2025). One model to rule them all? Choosing two-dimensional spaces for European political landscapes with VAA data. *Journal of Information Technology & Politics*. Advance online publication. <https://doi.org/10.1080/19331681.2025.2469166>
- Toshkov, D., & Romeijn, J. (2021). How to estimate the policy preferences of party supporters: Disaggregating data from voting advice applications versus modeling survey responses. *Electoral Studies*, 74, Article 102403. <https://doi.org/10.1016/j.electstud.2021.102403>
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15(4), 582–599. <https://doi.org/10.1017/S1755773923000103>
- Vogels, V. (2023, November 2023). Klimaatplannen van de politiek op een rijtje: doe de kieswijzer over het klimaat. *NPO Radio 1*. <https://www.nporadio1.nl/nieuws/binnenland/6250498b-45de-4bf6-839e-1230fc0f83da/klimaatplannen-van-de-politiek-op-een-rijtje-doe-de-kieswijzer-over-het-klimaat>
- Wagner, A. (2017). A micro perspective on political competition: Electoral availability in the European electorates. *Acta Politica*, 52(4), 502–520. <https://doi.org/10.1057/s41269-016-0028-7>
- 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>
- Wagner, S., Wurthmann, L. C., & Thomeczek, J. P. (2023). Bridging left and right? How Sahra Wagenknecht could change the German party landscape. *Politische Vierteljahresschrift*, 64(3), 621–636. <https://doi.org/10.1007/s11615-023-00481-3>
- Wheatley, J., Carman, C., Mendez, F., & Mitchell, J. (2014). The dimensionality of the Scottish political space: Results from an experiment on the 2011 Holyrood elections. *Party Politics*, 20(6), 864–878. <https://doi.org/10.1177/1354068812458614>
- Wurthmann, L. C., & Wagner, S. (2024). On the association of attitudes towards NATO and weapon deliveries for Ukraine with vote intentions in Germany. *German Politics*, 34(1), 1–24. <https://doi.org/10.1080/09644008.2024.2372562>

- Zhirnov, A., Thomeczek, J. P., Scotto Di Vettimo, M., López Ortega, A., Krouwel, A., Antonucci, L., Di Stefano, R., & Kersting, N. (2025). How congruent are populist parties with their constituencies? Evidence from the 2019 European parliament elections in Italy, the Netherlands, Germany and Sweden. *Electoral Studies*, 94, Article 102906. <https://doi.org/10.1016/j.electstud.2025.102906>
- Zicht, W., & Cantow, M. (2025). *Infratest dimap* [Data set]. Wahlrecht.de. <https://www.wahlrecht.de/umfragen/dimap.htm>

About the Author



Jan Philipp Thomeczek is a postdoctoral researcher at the University of Potsdam, chair for Comparative Politics. In his research, he focuses on voting behaviour, political parties, voting advice applications, and populism. He is currently a principal investigator for the DFG-funded EtaPPP project on the establishment of BSW.

Voting Advice Applications and Their Impact on Ideological and Affective Polarization

Maxime Walder ^{1,2} , Jan Fivaz ³ , and Daniel Schwarz ³ 

¹ Department of Political Science and International Relations, University of Geneva, Switzerland

² Department of Social Science, University of Basel, Switzerland

³ Center for Public Management, University of Bern, Switzerland

Correspondence: Maxime Walder (maxime.walder@unige.ch)

Submitted: 29 August 2025 **Accepted:** 24 November 2025 **Published:** 28 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Political polarization has been one of the most researched topics in political science over the past few years. While many scholars focus on the factors that foster ideological and affective polarization, less is known about the drivers of depolarization. In this article, we are addressing this gap by investigating the potential effects of using voting advice applications (VAAs) on both the ideological polarization of vote choice and the level of affective polarization among voters. Using data from the Swiss VAA Smartvote and from a field experiment within a two-wave panel survey conducted before and after the 2023 Swiss national election, this article presents partial evidence that using a VAA during election campaigns can contribute to a reduction in ideological and affective polarization among voters. Our findings have significant implications for the role of VAAs in increasingly polarized Western democracies.

Keywords

affective polarization; ideological polarization; vote choice; voting advice applications

1. Introduction

This article brings together two developments of modern democracies that are typically studied separately in research. On the one hand, the increase in ideological and affective polarization is well documented (e.g., Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017). On the other hand, the usage of voting advice applications (VAAs)—online platforms that allow voters to compare themselves with candidates or parties based on their positions on political issues—is increasing in liberal democracies (e.g., Garzia &

Marschall, 2014; Munzert & Ramirez-Ruiz, 2021). This research links these two phenomena by focusing on the potential impact of VAA use on the level of ideological and affective polarization of voters.

One key aspect of modern democracies is the growing ideological and affective polarization of both the public and political elites (Abramowitz & Saunders, 2008; McCarty, 2019; Poole & Rosenthal, 2000, 2001; Stonecash et al., 2003). While the United States is often the reference example, these dynamics are also observed across Western democracies (Boxell et al., 2020; Gidron et al., 2019; Wagner, 2021). These growing divisions have an impact on democratic norms, threaten the proper functioning of democratic institutions, and can lead to their gradual erosion (Gidron et al., 2020; Hajnal, 2025; Iyengar & Krupenkin, 2018; Kingzette et al., 2021). The ideological and affective polarization of political elites and the mass public are increasingly the subject of scientific research to better understand these fundamental dynamics.

In recent years, VAAs have established themselves as an indisputable tool for providing voters with information on candidates' positions on issues in the run-up to elections. A growing body of literature investigates the effect these tools have on turnout (Germann & Gemenis, 2019; Manavopoulos et al., 2018; Munzert & Ramirez-Ruiz, 2021), issue voting, and vote switching (Gemenis & Rosema, 2014; Germann et al., 2023; Tromborg & Albertsen, 2023). However, these tools also broaden voters' overview of the pluralistic nature of the political system. Voters who use VAAs learn how their position may align with candidates from various parties, sometimes including parties that they had previously not considered as potential vote options, which could affect their actual electoral behavior and also more fundamental political attitudes (Pianzola et al., 2019).

In this article, we aim to contribute to both the literature on VAAs and on political polarization by studying the effect that the use of these tools has on the polarization of vote choice as well as on the affective polarization of voters during election campaigns. The pluralistic recommendations that VAA users are exposed to reduce the perception they have of partisan sorting—their perceived alignment between their ideology and their partisanship—and affect the extremism of voters' decisions, thereby reducing the polarization of their vote choice. Furthermore, exposure to candidates they had originally not intended to vote for reduces out-group animus, which is arguably a significant driver of the increase in affective polarization levels (Druckman et al., 2021; Iyengar et al., 2012, 2019; Webster & Abramowitz, 2017). By studying the relationship between VAA use, ideological polarization, and affective polarization, we aim to provide an answer to the question: *How does the use of VAAs affect the polarization of vote choice and affective polarization of individual voters during election campaigns?*

We draw on findings from a field experiment conducted before the 2023 national elections in Switzerland. In the experiment, a randomly selected treatment group was informed about the possibility of using a VAA as an additional source of information ahead of the upcoming elections. Using a panel survey, we investigate how the use of VAAs by individual voters impacts the change in the polarization of their vote choice and their affective polarization compared to non-VAA users. We estimate both the effect of using the VAA and the causal effect of the encouragement treatment—the invitation to use the VAA—using the Complier Average Treatment Effect (Gerber & Green, 2012). In doing so, we enhance our understanding of whether and how such tools may (de-)polarize the public. Our results provide mixed evidence regarding a link between VAA use and vote polarization. While VAA users seem to decrease the polarization of their vote choice more than non-users, as well as increase their positive affective evaluation of parties, these effects are not consistent

across all parties. Furthermore, the causal effect of our treatment does not show a clear indication that these effects are robust. Our findings nevertheless indicate a non-null relationship between VAA use, ideological polarization, and affect for parties, providing, to our knowledge, the first evidence on the link between VAA use and polarization.

In the following sections, we first review the literature on VAAs and position our study within it. We then detail the theoretical mechanism linking VAA use and the polarization of voters' choice and its effects. Finally, we present and discuss our findings.

2. What Is the Effect of Voting Advice Applications?

Since its emergence over 20 years ago, a fairly large body of literature has explored VAAs as well as the effects of their use in various ways. A first strand of VAA research focuses on the sociodemographic characteristics of voters using VAAs. Early studies found that male, younger, well-educated, and politically interested individuals are disproportionately more likely to use VAAs (Hooghe & Teepe, 2007; Manavopoulos et al., 2018; Marschall, 2014; Marschall & Schultze, 2012, 2015). Walder et al. (2024) find that this remains largely true today, at least with regard to age, educational level, and political interest. These findings also align with the results of a longitudinal study on VAA use (Albertsen, 2022), which show that while the gap in age, education, and political interest between VAA users and non-users remains, the gender gap in VAA use has lost its significance in recent years.

Another set of studies on VAAs focuses on the methodological aspects related to the design of VAAs and their impact on various outcomes. Walgrave et al. (2009) study statement selection and show that it affects the recommendation of VAAs at the individual and aggregate levels. Similarly, Lefevere and Walgrave (2014) used Belgian VAA data and provided evidence that it had an impact on parties and voter matching. Additional studies have shown that statement selection is a key design choice influencing VAA outcomes (Louwerse & Rosema, 2014; Rosema et al., 2014). The literature also investigate how response scales and matching algorithms affect recommendations (Gemenis, 2024; Louwerse & Rosema, 2014; Otjes & Louwerse, 2014; Romero Moreno et al., 2022; Rosema & Louwerse, 2016). Recently, several studies have focused on the normative, ethical, and methodological limitations in VAA design of questionnaires, also discussing the implications of ethical concerns for stakeholders, as well as the use of AI to improve questionnaires (Gemenis, 2024; Stockinger et al., 2024). Finally, studies also discuss self-selection biases and their methodological implications for VAA research (Munzert et al., 2020; Pianzola, 2014) and examine matching techniques to address this bias (Gemenis & Rosema, 2014; Germann & Gemenis, 2019).

Finally, VAA literature has also focused its attention on the effects on voters. First, studies find positive but moderate evidence linking VAA use to the level of political knowledge (Heinsohn et al., 2016; Schultze, 2014). In their meta-analysis, Munzert and Ramirez-Ruiz (2021) found modest evidence for increased issue knowledge. Second, researchers have also investigated the relationship between VAA use and information-seeking behavior. Manavopoulos et al. (2018) found that VAA users are more likely to consume media related to politics, seek election-related information, and engage in political discussions. However, Mahéo (2017) found no statistically significant effect of VAA usage on attention to electoral campaigns or active election-related information seeking. Finally, researchers have studied the effect VAAs have on the political behavior of their users. Observational studies regularly find positive associations between VAA use

and higher turnout, especially among younger voters and the less politically engaged (Fivaz & Nadig, 2010; Gemenis & Rosema, 2014; Germann & Gemenis, 2019; Manavopoulos et al., 2018; Munzert & Ramirez-Ruiz, 2021). An exception in this regard is the study by Benesch et al. (2023), which finds no effect on turnout but increased ballot splitting. Experimental studies (Enyedi, 2016; Frese et al., 2024; Mahéo, 2017; Munzert et al., 2020) tend to find smaller or no significant effects on turnout, suggesting that observational studies might overestimate causal effects due to self-selection bias (Germann et al., 2023; Munzert & Ramirez-Ruiz, 2021). Furthermore, many observational studies report that VAA usage influences voters to align their vote choice with the VAA recommendations, thus increasing issue voting and vote switching (Garzia & Marschall, 2014; Gemenis & Rosema, 2014; Kleinnijenhuis et al., 2019; Ladner et al., 2012; Tromborg & Albertsen, 2023; Walgrave et al., 2008). However, field experiments with random assignment show mixed results regarding effects on vote choice: while Germann et al. (2023) and Pianzola et al. (2019) report significant effects on vote choice, others do not find them (Enyedi, 2016; Frese et al., 2024; Mahéo, 2017; Munzert et al., 2020).

Our article contributes to the literature on the effects of VAAs on their users by investigating a previously neglected dimension: the impact of VAA use on the ideological and affective polarization of voters. Although the primary aim of VAAs is not to decrease polarization but to help voters navigate a complex information environment, it is important to understand whether VAAs affect the political attitudes and behavior of voters beyond their main objectives.

3. Voting Advice Applications and Polarization

Political polarization has been a central focus of research in political science in recent years. Researchers have differed between ideological and affective polarization. Ideological polarization is generally defined as the extent to which issue positions and ideologies are opposed within a society (Fiorina et al., 2008; Hajnal, 2025; Iyengar et al., 2012; Lelkes, 2018). Affective polarization is rooted in the social identity theory (Tajfel & Turner, 1979), where partisanship functions as a social identity that divides the political world into an “us” (in-group) versus “them” (out-group) (Bornschieer et al., 2021, 2022; Zollinger, 2024). Affective polarization occurs when the distance between positive in-group affect and negative out-group affect rises (Iyengar et al., 2012, 2019; Wagner, 2021). Both forms of polarization have substantial consequences that threaten democratic accountability and the functioning of democratic institutions by fostering mistrust, decreasing political tolerance, and increasing partisan hostility (Gidron et al., 2019; Kingzette et al., 2021). Partisans may prioritize ideology over integrity, becoming more tolerant of undemocratic practices if they serve their ideological goals, which can erode democratic norms and lead to democratic backsliding (Gidron et al., 2020; Hajnal, 2025; Iyengar & Krupenkin, 2018).

Ideological and affective polarization are theoretically different concepts. Empirically, there is mixed evidence on their relationship. On the one hand, studies find that one can experience strong affective polarization even when ideological differences are modest or shrinking (Iyengar et al., 2012; Levendusky & Malhotra, 2016; Mason, 2015). On the other hand, studies have found that greater ideological distance leads to higher levels of affective polarization (Riera & Madariaga, 2023; Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017). Using longitudinal data, Lelkes (2018) finds a reciprocal relationship between ideological consistency and affective polarization, where each can reinforce the other over time. While affective polarization can be fueled by social identities, tribalism, and emotional group attachment beyond ideology, ideological consistency

and extremism also contribute to some extent to affective polarization (Druckman et al., 2021; Iyengar et al., 2019; Lelkes, 2018).

Beyond the relationship between ideological and affective polarization, several factors fuel polarization in general. First, the polarization of the political elites provides strong ideological cues to the electorate, reinforcing or shaping their ideological positions (Abramowitz & Saunders, 2008; Gidron et al., 2020; Rogowski & Sutherland, 2016). Individuals are motivated to adopt ideological positions consistent with their partisan identities due to cognitive biases and motivated reasoning, which reinforces ideological consistency and polarization over time (Lelkes, 2018). This increase in partisan sorting (Levendusky, 2009) strengthens partisan attachments and group-based distinctions (Iyengar et al., 2019; Mason, 2015). Second, the growing alignment between social and partisan identities reduces the number of cross-cutting identities, thereby lowering their moderating influence on polarization and facilitating stronger ideological divisions (Lelkes, 2018; Mason, 2015). Third, the growing partisan media system and the proliferation of online platforms and social media have intensified ideological divides by providing selectively biased information that reinforces existing ideological beliefs (Druckman et al., 2021; Gidron et al., 2020; Lelkes, 2018; Prior, 2013). In addition, negative political campaigns and strong media focus on polarization contribute to the increase in mass polarization (Iyengar et al., 2012; Sood & Iyengar, 2016), and evidence shows that campaigns generally make partisanship more salient (Michelitch & Utych, 2018).

Empirically, studies show a growing ideological polarization of political elites (Abramowitz & Saunders, 2008; McCarty, 2019; Poole & Rosenthal, 2000, 2001; Stonecash et al., 2003). For the polarization of the mass public, there is a conflict between a maximalist and a minimalist interpretation. The maximalists argue that ideological polarization among the mass public has increased significantly since the 1970s (e.g., Abramowitz & Saunders, 2008), while minimalists suggest that most of the electorate remains centrist, that the electorate's ideological distribution has not shifted dramatically, and that although partisanship and ideology are more tightly linked, this does not translate into a massive shift towards ideological extremes (e.g., Fiorina et al., 2008), which has been referred to as partisan sorting (Levendusky, 2009). Lelkes (2016) finds that, while the overall mass public has not experienced significant increases in ideological divergence or consistency, these measures have grown among strong partisans. This suggests that political polarization within the mass public may primarily result from increased sorting among strong partisans who align more closely with the polarizing positions of political elites. Consequently, elite polarization does not necessarily polarize the entire mass public but does so among strong partisans through this alignment process. Beyond their ideological polarization, studies highlight a growing affective polarization between party groups. Several pieces of research indicate that this increase in affective polarization is driven by growing animus towards the out-group (Iyengar et al., 2012), which now outweighs the in-group positive affect (Druckman et al., 2021; Iyengar & Krupenkin, 2018; Iyengar et al., 2019; Webster & Abramowitz, 2017). The increase in animus towards the out-party group increases with elite polarization (Rogowski & Sutherland, 2016), but this is not confined to the US, being widespread across Western democracies (Boxell et al., 2020; Gidron et al., 2019; Wagner, 2021).

Considering how elite polarization, partisan sorting, the disappearance of cross-cutting social identities, and the changes in the media environment of individuals affect polarization, how is the use of VAAs related to affective and ideological polarization? VAAs expose individuals to a variety of ideologies and issue positions of political parties and candidates. Exposure to varying issue positions and ideologies lowers the perception voters have of partisan sorting. VAA users are exposed to multiple parties and candidates' profiles that have

varying degrees of ideological closeness with them. This multiplication of choice creates a more balanced consideration in their decision-making process than they would otherwise experience during political campaigns (Pianzola et al., 2019), which are often more negative, vitriolic, and focused on the conflictual aspects of politics, with increasingly partisan media coverage that emphasizes the more polarizing aspects of the competition (Iyengar et al., 2019). Such effects are likely to occur particularly in countries with open-list electoral systems, such as Switzerland. Pianzola et al. (2019), based on an experimental study from Switzerland, were able to show that the use of a VAA increases the number of parties considered as viable voting options, and Benesch et al. (2023), another study based on Swiss data, found evidence that VAA use leads to an increase in vote-splitting. Since VAA users receive voting recommendations for candidates from different parties (see Figure A1 in the Supplementary File for an example), they are provided with a broader and more nuanced picture of available political alternatives and subsequently cast their votes for more parties compared to non-VAA users. Most parties are not positioned on the very extremes of the political space; thus, a larger number of parties receiving votes makes it much less likely that VAA users vote purely for extreme parties than voters who do not use a VAA. Thus, we expect that *VAA use decreases the polarization of vote choice* (Ideological Polarization Hypothesis).

Regarding affective polarization, while many studies have examined factors that increase polarization, few have also investigated those that reduce animosity towards the out-party. In their study, Ahler and Sood (2018) showed how correcting misperceptions about out-group supporters reduces animosity towards them. Similarly, Levendusky (2018a) showed that priming a common national identity reduces affective polarization. However, not all attempts yielded positive results, as Levendusky (2018b) found no evidence that priming partisan ambivalence lowers affective polarization. We argue that VAA use corrects the perception voters have of the closeness to the in-group and the distance towards the out-group. The exposure to candidates from different parties and the variety of plausible and reasonable options available to VAA users impact their partisan attachment and group-based distinction. As a result, VAA users challenge the perception voters have of their partisan identity, making them reconsider the pre-existing perceived division between their in-group and their out-groups. As a result, they may feel less strongly about their in-group and more positively about their out-group, thereby reducing the gap between their affect for in- and out-group parties. Most studies find that the increase in affective polarization is driven by an increase in out-party animus (e.g., Iyengar et al., 2012, 2019; Mason, 2015, 2018; Webster & Abramowitz, 2017). However, by being exposed to recommendations for candidates of different parties, the use of VAA potentially results in less out-party animus. As a result, we expect that *VAA use reduces animus toward parties* (Animus Hypothesis). As animus for parties decreases, and animus drives affective polarization, we expect that *VAA use decreases the levels of affective polarization* (Affective Polarization Hypothesis).

This article studies how VAA use affects ideological polarization of voting behavior and the level of affective polarization during the 2023 Swiss National election. However, most of the literature focuses on the US case. It is thus important to situate the Swiss case in terms of ideological and affective polarization to position our analyses in a comparative perspective. First, Switzerland operates under a multi-party system with a strong fragmentation and a proportional representation electoral system, which tends to lower ideological polarization compared to the two-party, majoritarian electoral system in the US (Gidron et al., 2020; Riera & Madariaga, 2023). Despite this, empirical evidence shows that ideological polarization has increased in Switzerland over the last decades, both in the population (Sciarini, 2023) and between parties (Walder, 2025). Furthermore, in comparative perspective, Switzerland exerts high levels of affective

polarization (Gidron et al., 2019). Finally, Riera and Madariaga (2023) find that Switzerland is an average case in terms of the effect that political extremism has on affective polarization. In sum, Switzerland has an electoral and a party system that favors less ideological and affective polarization. Still, it exerts a high level of polarization and the interrelation between affective and ideological forms of polarization is average in a comparative perspective.

4. Data and Methods

We conducted a two-wave panel survey during the 2023 national election campaign in Switzerland. The first pre-electoral survey wave was conducted six weeks before the election and the second post-electoral survey wave was conducted right after the election, which was held on October 22, 2023. National elections in Switzerland are a good case to study for two main reasons. First, the level of polarization and the affective position, as well as the relation between the two concepts, is average compared to other Western democracies across Europe (Riera & Madariaga, 2023). Second, VAAs are widely used during election campaigns, and the Smartvote VAA has been active in every National election since 2003 (see www.smartvote.ch for more information). The VAA used during the National elections contained 75 questions and 84% of all candidates for the national election responded to it (4,999 candidates out of 5,925). At the end of the VAA, voters access a list of candidates ranked from the highest to the lowest match. The first page of voting recommendations contains the first 50 recommendations, and voters are then free to navigate, look at all the matches, consult matching candidates in other electoral districts, and click on individual candidates to have more information on their profiles. On the first page, the first names, last names, birthdates, political party, and matching scores of the candidates are visible to VAA users. Figure A1 in the Supplementary File presents a screenshot of the top of the first recommendation page. In this context, the VAA helps voters navigate this highly complex information environment.

Between the two survey waves, we invited a random sample of respondents from the first survey wave to use the Smartvote VAA. The invitation to use the VAA was sent by email to the first wave survey respondents on 22 September 2023 (see Appendix A of the Supplementary File for the full invitation text). This was before voters in any electoral district had received their mail-in ballot material, which arrived, depending on the canton, between 24 September and 1 October. A reminder email was sent on 6 October. While the reminder email was sent after voters had received their mail-in ballot material, analyses of the cumulative Swiss Election Study (Selects, 2025) indicate that during elections between 1971 and 2023, only 2 to 9% of voters cast their ballot directly when receiving the mail-in ballot material (see Table D1 in the Supplementary File). Table A1 in the Supplementary File presents the full overview of the timeline of the study. Overall, 3,941 respondents participated in the first survey wave and 3,290 replied to the second wave. The survey was designed on Qualtrics, survey respondents were recruited by Bilendi, and to ensure some representativeness of the sample, we included quotas on age, gender, and education levels. We removed respondents who had not answered at least six of the policy questions to compute their ideological position (see Section 5), which leaves 2,573 observations. We invited 1,727 of them to use the VAA between survey waves, and among them, 878 actually did so. Furthermore, as the VAA is a tool used by roughly 550,000 voters during the electoral campaign (based on our own calculation with the Smartvote data), we asked respondents in the control group whether they had used a VAA during the political campaign, and 21% of the 846 respondents in the control group reported having used it. In this article, we use both the encouragement treatment—the invitation to use the VAA—and the observed use as

independent variables to estimate how the use of VAA and the causal effect of our treatment affect the polarization of vote choice and the affective polarization of voters during the 2023 Swiss National election campaign. While the observed use compares VAA users and non-users, to estimate the intention to treat effect, we estimate the Complier Average Causal Effect (CACE) as formulated by Gerber and Green (2012).

In the pre-electoral wave, we asked respondents about their vote intention and their attitude towards the six main Swiss parties. We replicated these measures by asking respondents about their vote choice and affect for the six main parties in the post-electoral survey wave. Although Switzerland has a highly fragmented party system, with many parties presenting candidates in the national election, the six main parties in the current legislature are Die Mitte (centrist party), the FDP (liberal party), the GLP (green liberal party), the GPS (green party), the SP (social-democratic party), and the SVP (national-conservative party). These parties hold 193 out of the 200 seats and more than 90% of the national vote share. Our analyses focus on these six parties and analyze how the variation between vote intention and vote choice, as well as the difference between pre- and post-electoral affect for the parties, is influenced by VAA use.

5. Measuring the Polarization of Vote Choice

To measure the polarization of voting behavior during the 2023 national election, we relied on survey respondents' vote intention and vote choice, as well as candidates' and voters' responses to the VAA. In the first step, we ran a Bayesian Item-Response Theory (IRT) model to estimate the position of voters and political candidates in the same space. IRT models have been used in recent years to measure the ideological position of voters, groups of voters (Berwick & Caughey, 2025; Caughey et al., 2019; Cavaillé & Trump, 2015; Traber et al., 2023), or both political elites and voters in the same political space (Leimgruber et al., 2010; Masket & Noel, 2012; Walder, 2025). In this article, we apply the strategy of Walder (2024) and position both political candidates and voters on the same ideological dimension, combining data from VAA and survey data.

For political candidates and voters who used the VAA, we collected the responses they gave to the 57 questions in the VAA that used a the four point scale. These 57 items comprise policy-related questions covering a wide range of topics, including immigration, social welfare, economic policies, security, the environment, and others. For survey respondents who did not use the VAA, we asked six out of the 57 policy questions in the post-electoral survey, chosen to cover various topics and issue dimensions. Although the position measured for non-VAA users only considers a small sample of the complete set of policy questions, research by Walder (2024) has shown that using a handful of similar questions enables an accurate measure of the latent position of non-VAA users. The policy questions are categorical with four response categories. We thus use a 2-parameter ordered IRT model that can be formalized as:

$$P_r(Y_{ij} = k | \theta_j, \beta_i, \alpha_i) = \phi(\theta_j \beta_i - \alpha_{i,k-1}) - \phi(\theta_j \beta_i - \alpha_{i,k})$$

Where ϕ is the cumulative normal distribution, θ_j is the ideal position of actor j , $\alpha_{i,k}$ is the difficulty parameter of item i for the category k , and β_i is the discrimination parameter of item i . The term $P_r(Y_{ij} = k)$ is the probability of the actor j answering the k category on item i . The parameters $\theta_{i,k}$ are ordered so that the difficulty for category $k - 1$ ($\alpha_{i,k-1}$) is lower than the difficulty for the category k of the same item ($\alpha_{i,k}$). To measure the ideological position of vote choice, the main parameter of interest in the model is θ_j , as it indicates the position of actor j in the latent space. Figure 1 presents the distribution of ideological positions of candidates from the six parties we consider to measure vote choice ideology, as well as those of the voters.

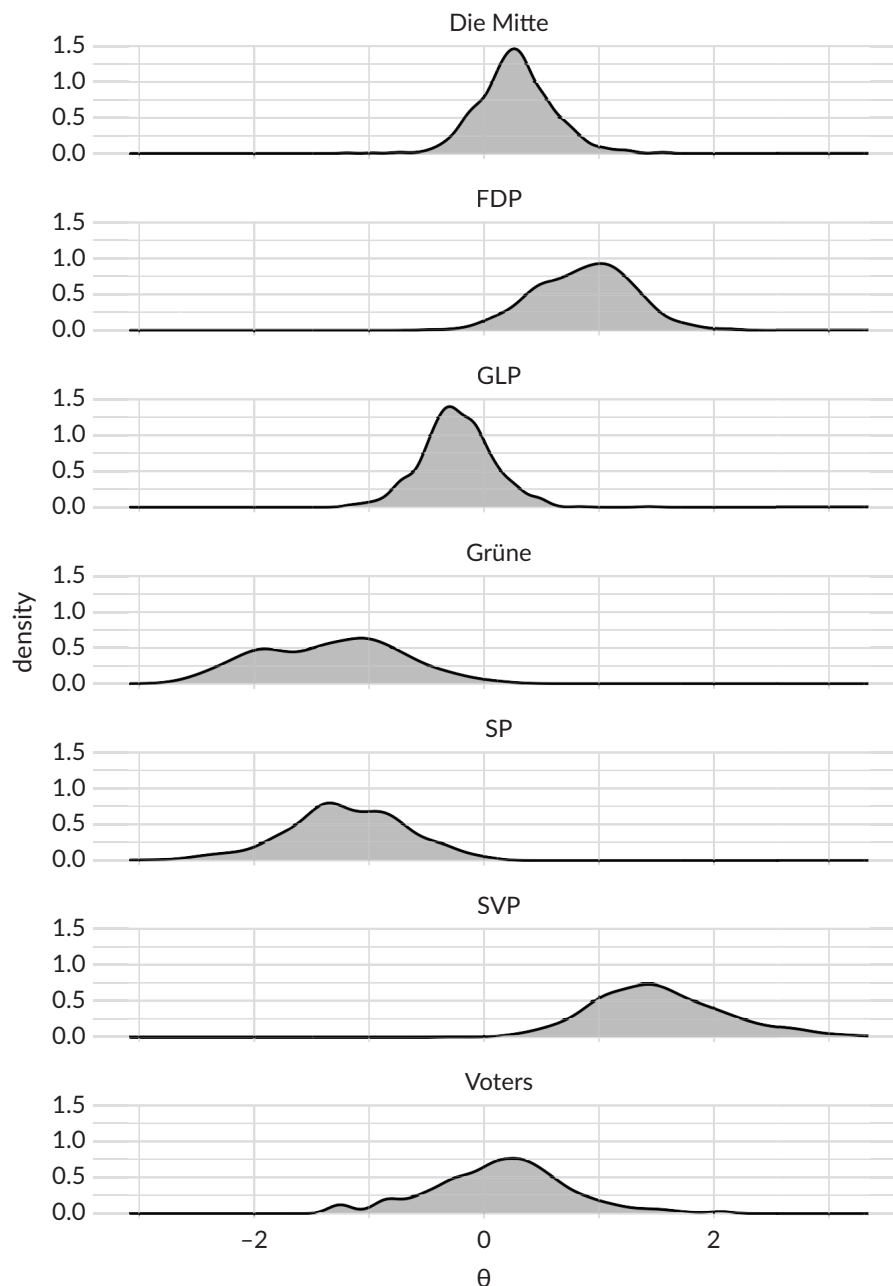


Figure 1. Distribution of latent ideological position for the political parties and for voters.

Figure 1 shows that voters' ideological positions are distributed on both sides of the status quo. We observe that the GPS and SP parties are positioned most on the left side of the political spectrum. In contrast, the distribution of SVP candidates' ideological positions is further to the right. Once the positions of the voters and the candidates were computed, we aggregated the positions of party candidates by party. We considered the mean as the average ideological position of the party. We then operationalize the ideological position of vote choice as the average position of the party respondents voted for the most.

Once we computed the ideological position of vote intention and vote choice, we measured the extremism of the vote choice with the absolute value of this ideological position. Positions closer to 0 indicate vote

intention and vote choice for parties closer to the status quo; the higher the value, the more extreme the vote intention and vote choice of respondents. Finally, we measure the polarization of vote choice with the difference between the extremism of vote choice and the extremism of vote intention. Values greater than 0 indicate a more extreme vote choice than the vote intention, and thus a more polarized vote choice than the vote intention. Values lower than 0 indicate a less extreme vote intention than the vote choice, and thus a more centralized position. Figure 2 presents the distribution of the polarization of the vote choice indicator.

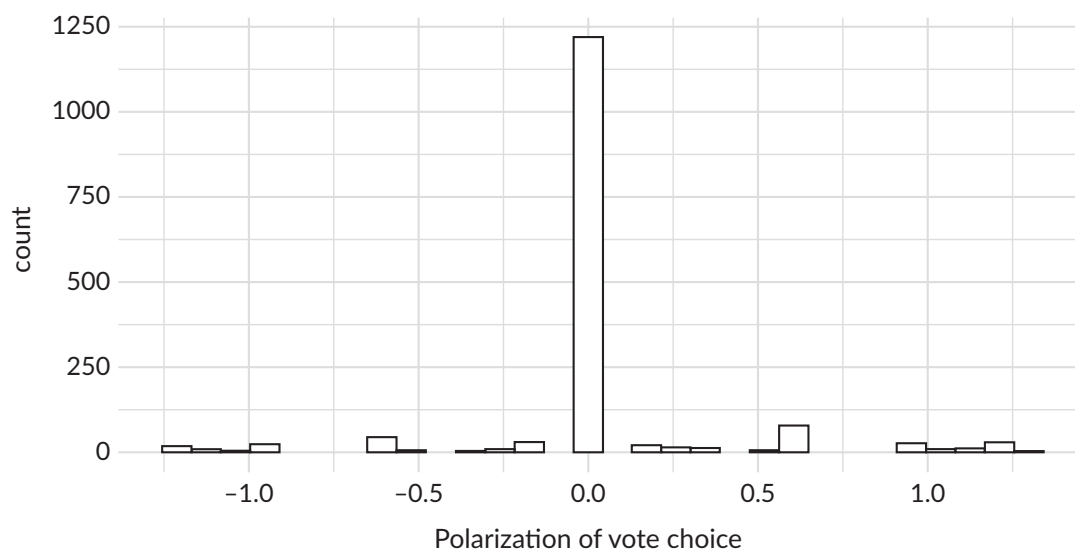


Figure 2. Distribution of the vote polarization indicator.

Figure 2 shows that individual voters are mostly voting for the same party that they intended to do in the early campaign. Indeed, for most voters, the difference in vote polarization is equal to 0, meaning that they voted for the same party they intended to. However, we observe that some voters changed the level of extremism in their vote choice compared to their initial intention. In the first part of the results, we use the difference in the extremism of vote choice as the dependent variable, indicating the polarization of vote choice during the electoral campaign.

5.1. Measures for Changes in Affective Polarization

To measure affective polarization, we rely on the indicators developed by Wagner (2021), who proposes two measures: mean distance and spread-of-scores. The key difference between these two indicators is that while the spread-of-scores measure computes the difference between the affect score of each party and the average affect score, the distance measure computes the difference between the affect score for each party and the maximum affect value for each respondent. Furthermore, to account for the fact that not all parties have the same weight in the electoral competition, Wagner (2021) also suggests using a weighted measure of affective polarization, which, instead of taking into account the average of these differences, multiplies them by the vote share of each party. He also emphasizes that the spread measure is better suited to measure affective polarization in multiparty systems. He nevertheless shows that these measures are correlated with one another. We thus compute the weighted and unweighted spread-of-scores and mean distance, and replicate the analyses for each of these indicators.

To estimate voters' affective relation for each of the six main parties, we provided them with an 11-point scale. We asked them to evaluate their affect for each party with the following question: "What is your opinion of the following national political parties? Please rate your answer on a scale of 0 to 10, where 0 means you dislike the party very much, 10 means you like it very much, and 5 is between the two categories." We asked respondents this question in both the first and second waves, which provided individual-level measures of affective polarization before and after the election. Figure 3 presents the distribution of the weighted and unweighted spread-of-scores and mean distance measure of affective polarization before and after the 2023 Swiss national election.

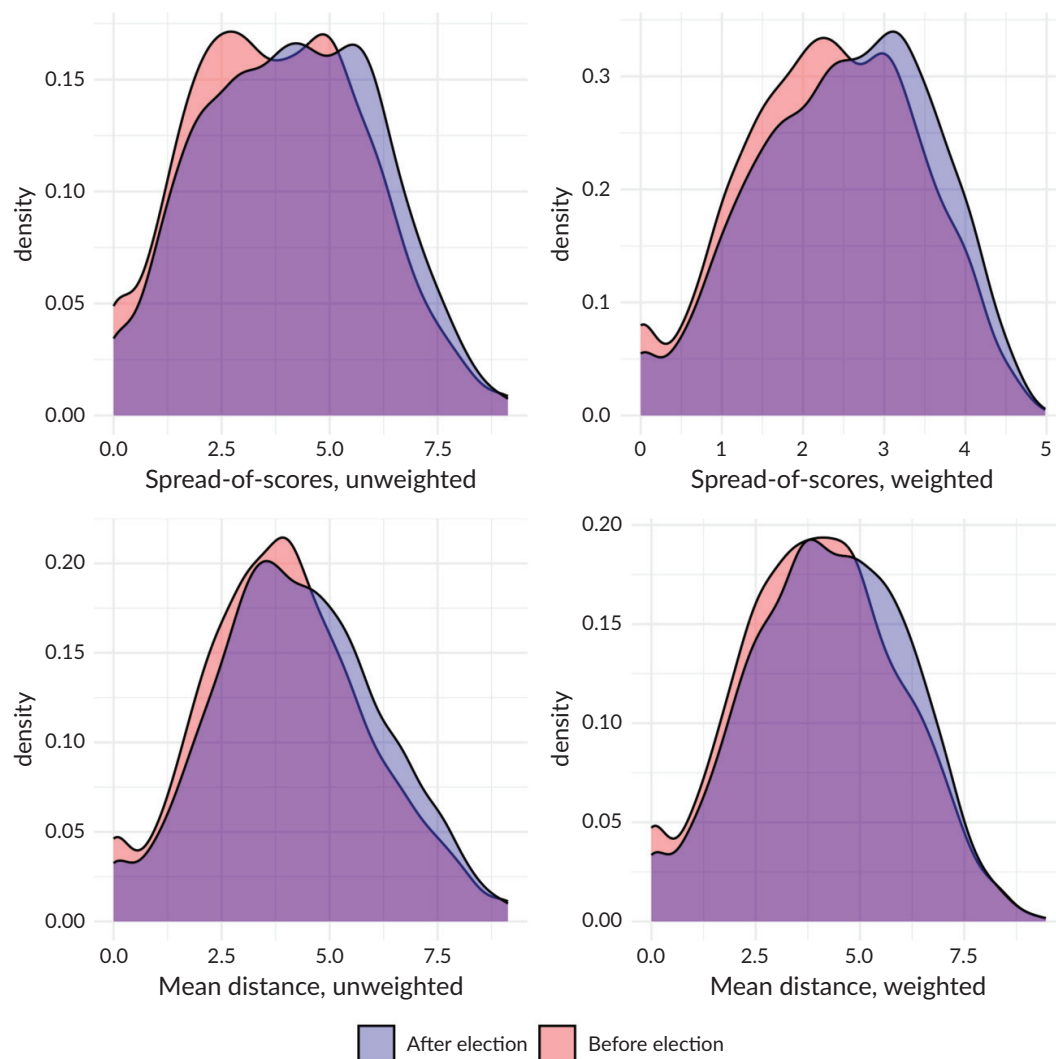


Figure 3. Distribution of affective polarization in the pre- and the post-electoral survey waves.

Figure 3 shows the affective polarization levels before and after the election campaign with all the different metrics developed by Wagner (2021). The four panels of this figure present the distribution of affective polarization scores before and after the election campaign using the four metrics developed by Wagner (2021). Figure 3 also shows that the post-electoral distribution of ideological polarization is more concentrated toward higher scores after the election than before. This result is consistent with the finding of Sood and Iyengar (2016), who found that affective polarization is substantially higher on election day than a

year before the election. In our analyses, we compute the difference in individual affective polarization before and after the election.

5.2. Measurement Models

Our study employs an encouragement treatment that invites voters to use the Smartvote VAA, which is also available to voters outside of the study. Thus, in addition to non-compliers in the treatment group (respondents who were invited to use the VAA but did not), there are also non-compliers in the treatment group (respondents who were not invited to use the VAA but did so anyway). In the post-electoral survey wave, we asked respondents to indicate whether they used a VAA during the political campaign or not. We can thus estimate the share of the treatment and control groups that used the tool during the political campaign. Overall, 51% of respondents in the treatment group used the VAA, compared to 21% of VAA users in the control group.

We conduct two types of analyses. The first type of analysis aims to estimate the effect of using the VAA on the polarization of vote choice, the affective relations voters have on the different parties, and their level of affective polarization. These analyses provide an overview of the general effect of VAA use on the various dependent variables considered in this article. This first set of analyses compares the VAA users to the non-users and estimates the effect that using the VAA has on the polarization of vote choice and affective polarization of voters. In this article, we label these estimates the effect of “observed use” of the VAA. Second, we leverage the random treatment allocation (the invitation to use the VAA) to compute the CACE as formulated in Gerber and Green (2012). Using a two-stage least squares regression with the treatment assignment as the instrumental variable and the observed treatment as the independent variable, the models estimate the causal effect of our treatment on the different dependent variables—the effect the invitation to use the VAA has on vote polarization and affective polarization. These two estimations present different effects that are interesting to study, as they do not necessarily tell the same story about the use of VAA. While the direct model estimates the direct effect of using the tool, the CACE estimates the causal effect of our treatment—the effect the tool has on respondents who would not have used it if they had not been in the treatment group.

We run these two measurement models on different dependent variables: the polarization of vote choice, the difference in affect for parties, and the difference in affective polarization. For each of these dependent variables, we control for the level of political interest of voters, their age, and their gender, as these characteristics influence the probability of using the VAA and complying with the treatment (Walder et al., 2024). As shown in Tables C1 and C2 in the Supplementary File, although the invitation to use the VAA is randomized, the share of VAA users in the treatment and the control group varies substantially by levels of political interest and for the different age groups. Furthermore, as we analyze the difference between post-electoral and pre-electoral values, these values are constrained by their pre-electoral value. Indeed, for instance, if voters intend to vote for the most extreme party in the pre-electoral survey wave, the polarization of the vote choice indicator can only have values that are negative or equal to 0. On the contrary, if voters intend to vote for the most centrist party, the polarization of their vote choice can only be null or positive. This logic also applies to the difference in the level of affect between parties and the level of affective polarization. Thus, we control for vote intention extremism in our analyses on vote polarization, we control for the initial level of affect in our analyses on the difference in affect between parties, and we

control for pre-electoral affective polarization when analyzing the effect VAA use has on affective polarization change. Additional information on key socio-demographic variables—question wording and descriptive statistics—is presented in Appendix B in the Supplementary File.

6. Results

In this section, we present the results of our analyses. First, we present descriptive evidence showing that VAA users are exposed to recommendations for candidates from different parties and varying shares of in-party candidates. Second, we present the analyses on the effect VAA has on the polarization of vote choice. Third, we present results on the effect VAA has on affect for the different parties, and finally, how it impacts the affective polarization of voters.

6.1. Descriptive Statistics: The Multi-Partisan VAA Recommendations

In the first part of the results, we present descriptive evidence to support our claim that VAA users experience high levels of exposure to various party candidates, which serves as the starting point of our theoretical argument leading to the development of our hypotheses related to the effect of VAA use on affective and ideological polarization. Figure 4 presents the distribution of the number of parties VAA users are exposed to and the share of in-party candidates that are recommended to them in the first 20 matches.

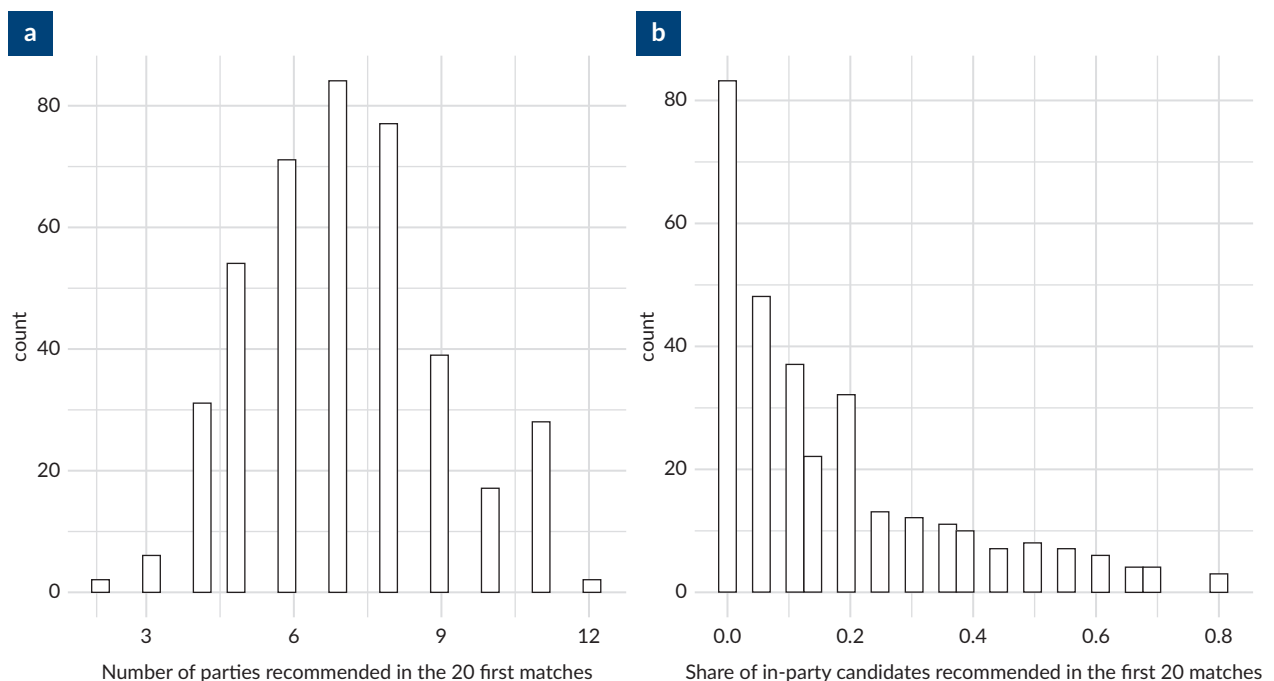


Figure 4. Number of parties (a) and share of in-party candidates (b) present in the first 20 matches of surveyed VAA users.

In Figure 4a, we see that the number of parties to which VAA users are exposed at the top of the recommendation list (the first 20 matches) varies between 1 and 12 and follows a relatively normal distribution, slightly skewed on the left, with a mean of 7 parties. Additionally, we observe that a small percentage of VAA users have only one or two parties in their recommendation list. In Figure 4b, we plot the

share of in-party candidates in the recommendation list, which means candidates from parties voters intend to vote for. We observe that a large majority of users see no, or only a small share, of in-party candidates in the first 20 recommendations, and only a small portion have a majority (more than 50%) of in-party candidates. This result shows that VAA users experience great exposure to out-group party candidates, which supports our claim that using VAAs challenges the perception voters have of partisan sorting, and their in-group and out-group affect.

6.2. VAA Use and Vote Polarization

In this section, we present the results of our analyses on the effect of VAA use and the causal effect of the treatment on the polarization of vote choice. Because voters are subjected to multiple parties in their recommendation list, this lowers the perception of partisan sorting, which decreases the polarization of their vote choice. Figure 5 presents the causal effect of our treatment and the effect of using VAA on the polarization of vote choice.

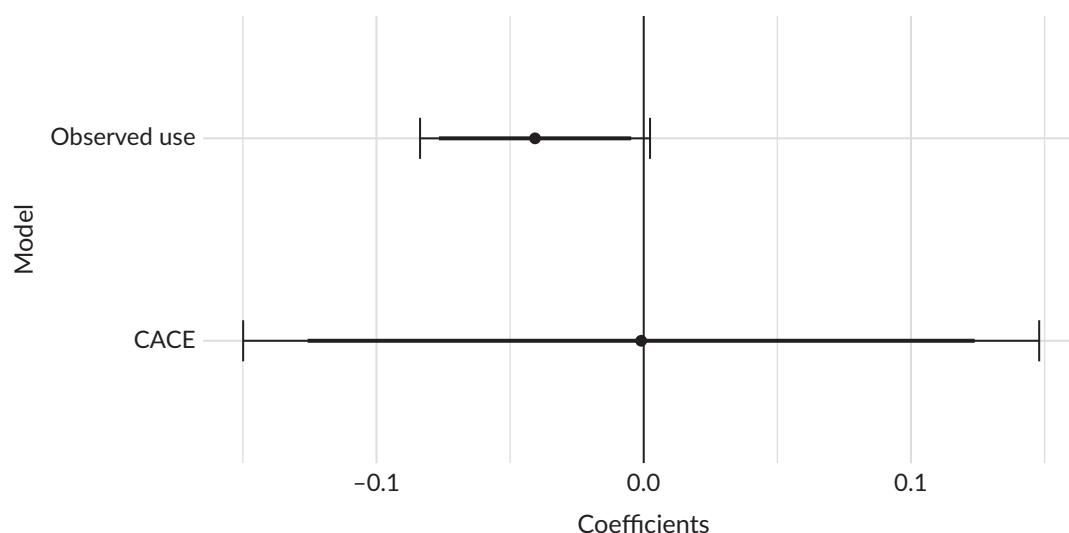


Figure 5. Effect of VAA use on the polarization of vote choice.

Figure 5 shows no significant effect of VAA use or of our treatment on the polarization of vote choice. However, we see that the VAA use has a negative effect on the polarization of vote choice, with a p -value close to the standard threshold of confidence ($p = 0.06$). While this is not conclusive evidence, our results tend to indicate that if VAA has an effect on the polarization of vote choice, it is more directed towards a depolarizing one. However, when examining the causal treatment effect, we see that there is no clear effect of the treatment on the polarization of vote choice. Although using VAA seems to decrease the polarization of vote choice, our results do not provide sufficient evidence to confirm the ideological polarization hypothesis.

6.3. VAA Use, Party Affect, and Affective Polarization

In this section, we present the results on the change in affect voters have on parties and the change in their level of affective polarization. Figure 6 presents the effect of VAA use and the causal treatment effect on the difference in affect for parties between the pre- and the post-electoral survey waves.

The left side of Figure 6 shows that the treatment has no causal impact on the levels of affect for the different parties. Indeed, in no instance is the effect on parties significantly affected by our treatment. While this suggests that VAA has no clear impact on the change in affect voters have on parties, the right side of Figure 6 tells another story. Indeed, we see that for half of the parties, the use of VAA significantly increases their affective evaluation by voters. This effect is significant for the Mitte party and the GLP, which are the two centrist parties. However, the effect is not significant for the SP and the FDP. Furthermore, while the GPS is not a centrist party, VAA use seems to increase the affective evaluation of this party by voters. Finally, we see that the affect for the SVP seems to decrease with VAA use, which goes against the animus hypothesis.

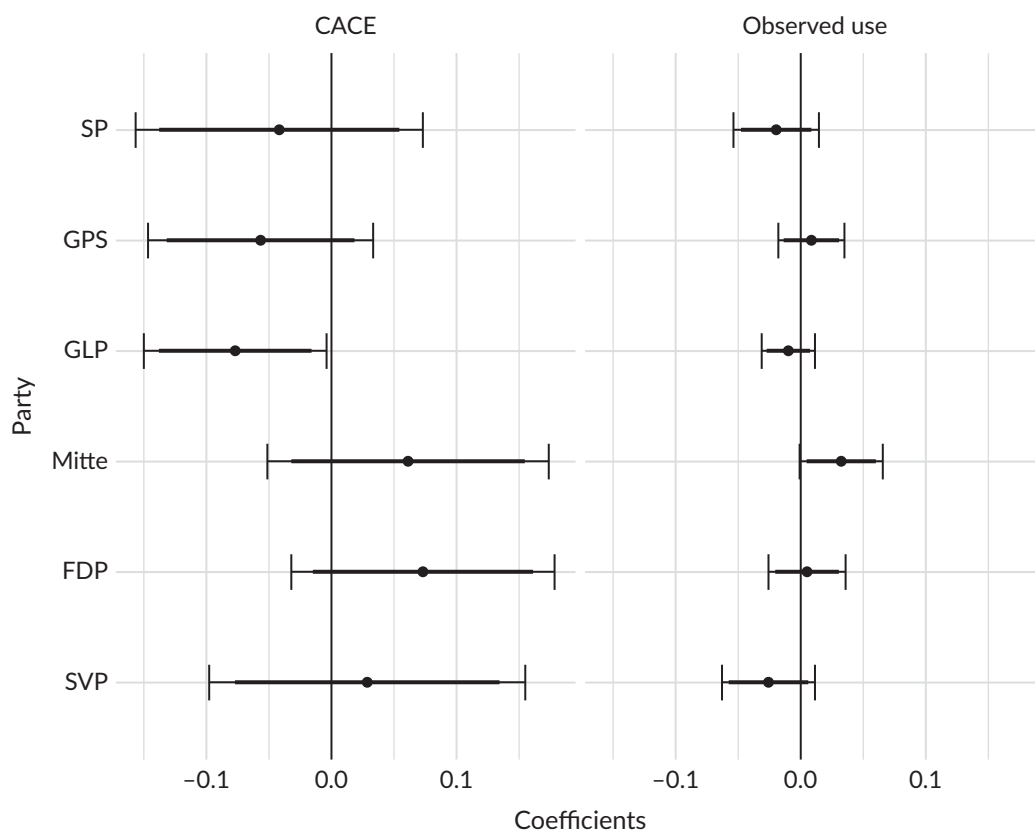


Figure 6. Effect of VAA use on the difference in affect for the different parties.

Figure 6 shows no clear trend in the effect VAA has on party effect. On the one hand, the causal effect of the treatment has no clear impact on the level of affect for parties. On the other hand, while the effect of using VAA is positive for most parties, it is negative for the SVP and the GLP, and none of the estimates reach the 95% threshold of significance.

The final part of our analysis focuses on the causal effect of the treatment and VAA use on individual change in affective polarization. Figure 7 presents the results of these analyses.

Figure 7 shows no clear effect of VAA use or treatment effect on the change in affective polarization. Moreover, the causal effect of the treatment and the effect of using VAAs seem to go in different directions. Indeed, while the left side of Figure 7 shows a non-significant positive effect of the treatment, indicating an

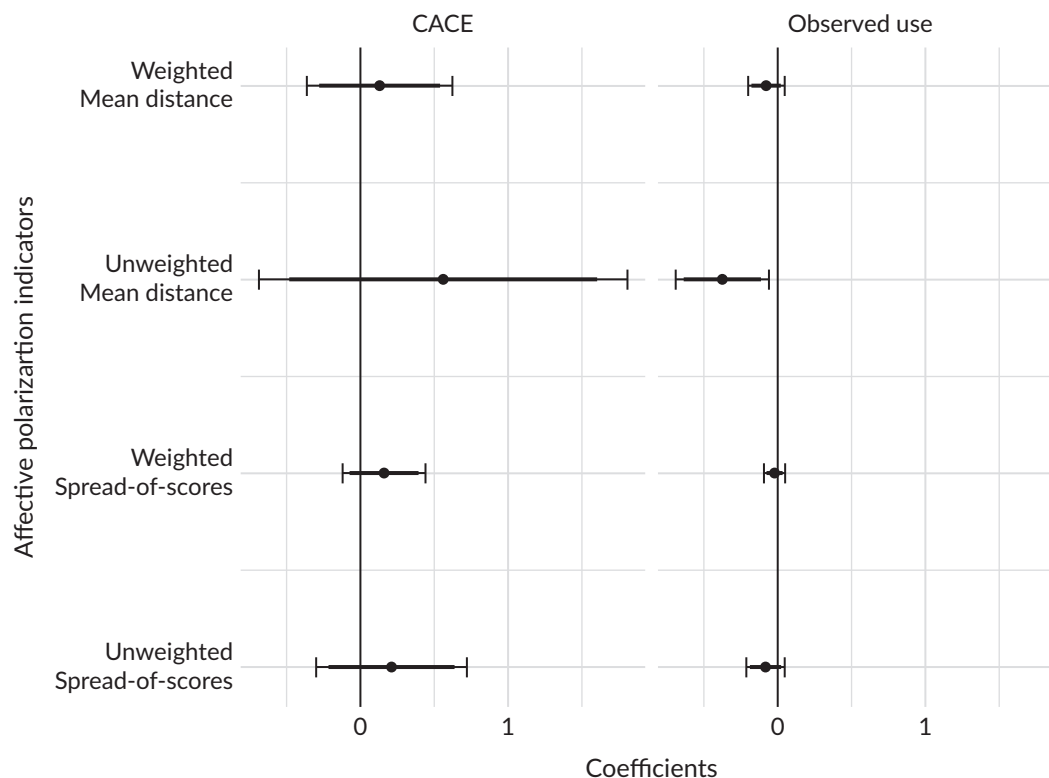


Figure 7. Effect of VAA use on the difference in affective polarization.

increase in affective polarization, the right side, focusing on the effect of VAA use, systematically presents negative estimates, which are not significant except for the unweighted distance indicator. This final finding does not enable us to confirm the affective polarization hypothesis.

7. Discussion

The results indicate no clear effect of VAA or treatment effect on the polarization of vote choice, the level of affect for parties, and the levels of affective polarization. However, this does not mean that VAAs generally have no impact on voter polarization. Indeed, if we focus our attention on the observed use of VAAs, most results go in the expected direction and indicate decreasing polarization levels. Figure 5 shows a negative effect of VAA use on the polarization of vote choice, which does not pass the 0.05 significance threshold generally accepted as significant, but is quite close to it. VAA use also has a positive impact on the change in affect for five out of six parties, but is not significant for the SP and the FDP, and is negative for the SVP. Finally, the use of VAAs has a consistent negative effect on the level of affective polarization, but again fails to pass the significance test. Taken all together, these results are encouraging but not strong enough to have a definitive conclusion on the use of VAAs.

Furthermore, the CACE hinders our confidence in these results. Indeed, none of the results are significant, and the CACE even indicates an opposite effect to that expected by the hypotheses, especially for the affective polarization hypothesis. While it is difficult to understand precisely what could explain such small effects and why the CACE and VAA use show such different findings, some elements related to VAA use and

the distribution of the treatment should be taken into account. Section C of the Supplementary File provides empirical evidence that the compliance with the treatment and the control group is not random and is influenced by pre-treatment characteristics.

Importantly, VAA users are not randomly distributed among voters and exhibit higher levels of political interest, education, and seek more information from politically relevant sources (Albertsen, 2022; Hooghe & Teepe, 2007; Manavopoulos et al., 2018; Marschall, 2014; Marschall & Schultze, 2012, 2015). Walder et al. (2024) show that perfect knowledge about the availability of a VAA does not correct for these biases in the self-selection of voters with higher interest. However, they highlight how, under perfect information conditions, a VAA would be used by a greater number of people with lower digital literacy skills, i.e., older voters. These observations suggest that not only are VAA users distinct from other voters, but also that the invitation to use a VAA has a differential impact on voters. Table C1 in the Supplementary File shows that the difference in VAA use between voters with different levels of political interest is similar in the treatment and the control group. This means that while the treatment encourages voters to use the tool, it does not mitigate the difference in usage between voters with different levels of political interest.

Additionally, two elements also warrant further attention. First, to estimate the effect of VAA use on vote polarization, vote intentions for the different parties must be similarly distributed across the compliers and non-compliers of the treatment and control groups. Figure C1 in the Supplementary File presents the distribution of party vote share for the different treatment groups. In Figure C1 in the Supplementary File, we observe that both the treatment group non-compliers and the control group compliers have higher vote shares for the conservative party SVP, the most extreme party according to our calculations. Furthermore, they also have the lowest vote share for the GLP, which is one of the most centrist parties in the country (see Figure 1). This difference in the distribution of vote intention among parties with different ideologies impacts the distance from the status quo of the vote intention of VAA users and non-users. Table C3 in the Supplementary File shows that while non-VAA users intend to vote for a party that is, on average, above 1 point away from the status quo, the VAA users intend to vote for parties that are, on average, below one point away from the status quo. The uneven distribution of this essential pre-treatment variable poses a problem in estimating the effect of VAA use and the CACE on vote polarization. Indeed, as non-VAA users exhibit more extreme vote and less centrist vote intentions, they also have more room for depolarization than VAA users.

Second, voters in the different groups should also show similar distributions of pre-treatment affective polarization levels. Table C4 in the Supplementary File presents the average affective polarization score for VAA users and non-users in the treatment, the control group, and the whole sample. This table also shows that VAA users have, on average, lower levels of affective polarization than VAA users before the treatment. These differences make it difficult to estimate a reliable causal effect of the treatment when computing the CACE. They may explain the differing directions of the results between the effect of VAA use and the CACE on the effect for parties, as well as the level of affective polarization.

Overall, VAA users have less extreme vote intentions, lower levels of pre-treatment affective polarization, and higher levels of political interest. The additional analyses in Appendix C of the Supplementary File show that these differences remain in the treatment group. Indeed, Table C5 in the Supplementary File presents the results of a regression model in which the extremism of vote intention has a significant effect on the probability

of using the VAA in the treatment group, the control group, and the whole sample. Despite the lower room for depolarization in the group that uses the VAA, our results indicate that the VAA has a depolarizing effect on vote choice.

8. Conclusion

To our knowledge, this article is the first study ever to investigate the effect that VAAs have on voter polarization. We study the impact of VAAs on the polarization of vote choice, the affect voters have for parties, and their levels of affective polarization, using a field experiment conducted within a two-way panel survey wave during a national election campaign. We argue that VAAs, by exposing voters to candidates from a wide range of parties, lower the perception of partisan sorting and challenge voters' partisan identity by recommending closely congruent out-party candidates. While our results do not provide strong evidence on the relationship between VAA use and ideological or affective polarization, the evidence presented points to a decrease in vote polarization and a decrease in party animus, albeit not for all parties.

By investigating the relationship between VAAs and polarization, this article contributes to the literature on depolarization. The evidence presented suggests a depolarizing effect, although, as highlighted, the non-random nature of VAA use, even with an encouragement treatment, makes it challenging to form definitive conclusions. While many researchers study what affects and drives polarization, only a few studies investigate elements that have the potential for depolarization (Iyengar et al., 2019). Both ideological and affective polarization threaten democratic accountability, increasing mistrust in political institutions, decreasing tolerance towards one another, and fostering hostility (Gidron et al., 2019; Kingzette et al., 2021). Growing polarization and negative partisanship threaten democratic norms by prioritizing ideological goals, even through undemocratic practices, which represents a rapid slope toward democratic backsliding (Hajnal, 2025; Iyengar & Krupenkin, 2018). Given the threats posed by growing polarization, future studies should aim to better understand how initiatives, such as VAAs but not limited to them, can lower polarization levels.

While our study addresses questions regarding the link between VAA use and ideological or affective polarization, several aspects remain underexplored. First, while we study how VAA use generally affects the levels of voters' polarization, we only compare voters who use VAA and voters who do not use the tool. However, the effect the VAA has among VAA users may vary depending on the recommendations voters get from the VAA. Indeed, while centrist voters who intend to vote for a rather extreme party might experience a depolarization of their vote choice following its use, the opposite is true for a rather extreme voter who intends to vote for a rather centrist party. Future studies should investigate the moderating impact that vote recommendations from VAA have on the polarization of vote choice and the affective polarization of voters. Second, in Switzerland, the vote is an open list system, meaning that voters can add candidates from other parties to their list. Our study focuses on the change in party list between the vote intention and the vote choice, but other aspects of voting (i.e., adding candidates from another party to the list) may also be affected by the use of VAA. While the cumulative Swiss Electoral studies indicate that the share of voters who do so is relatively limited (between 13% and 30% depending on election year, see Table D2 in the Supplementary File for a full overview), the use of VAA may also affect the addition of out-party candidates to the party list. Finally, the number of candidates and available seats in the 26 cantons varies between 2 for small cantons and 35 for the canton of Zurich. This affects both the number of candidates and the number of choices voters have to make. The number of seats in the district heavily impacts the number of candidates

and, thus, the potential for vote switching and the relevance of using a VAA. Indeed, VAA lose their importance if voters have to select one out of two candidates, or if they have to select thirty out of a thousand candidates.

We present evidence that indicates that there is a potential in VAAs to contribute towards a decrease or at least a moderation of political polarization. Previous VAA research has highlighted that VAA design and statement selection matter with regard to the impact a VAA can have on vote choice and various other aspects of voters' behavior (Gemenis, 2024; Louwerse & Rosema, 2014; Otjes & Louwerse, 2014; Romero Moreno et al., 2022; Rosema & Louwerse, 2016; Stockinger et al., 2024; Walgrave et al., 2009). However, the role the design of a VAA plays regarding the impact on their users' polarization levels is an unexplored dimension of VAA research. Future studies should examine how VAA designs and also communication measures can contribute to a better exploitation of this potential for reducing political polarization.

In general, the main goal of VAAs is to allow voters to make informed electoral decisions. However, VAAs potentially have additional benefits. They provide voters with a broad range of views on different parties or candidates, which can be seen as a valuable effect in its own right—especially in the case of Western democracies, where the public is more divided and polarized than ever. They carry the potential to balance ideas and highlight common political ground, while providing voters with a strong perspective to navigate complex information environments. These functions are crucial for the democratic institutions and socio-political cohesion of the public. This effect should not be underestimated, as democratic backsliding is gaining traction in more and more established democracies.

Acknowledgments

This work was supported by a grant from the Swiss National Foundation awarded to Prof. Nathalie Giger (Department of Political Science and International Relations, University of Geneva). We are grateful for her support in the making of this article.

Funding

This research was conducted as part of a project funded by the Swiss National Science Foundation (SNSF) [grant number: 197439]. Additional funding was also received by the University of Bern. Publication of this article in open access was made possible through the institutional membership agreement between the University of Geneva and the University of Basel with Cogitatio Press.

Conflict of Interests

Daniel Schwarz and Jan Fivaz are co-founders of Smartvote, the voting advice application that was part of the field experiment used in this article.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical restrictions.

LLMs Disclosure

The authors have used Grammarly to enhance the text of the article and Coral AI to conduct a comprehensive literature review.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, 70(2), 542–555.
- Ahler, D. J., & Sood, G. (2018). The parties in our heads: Misperceptions about party composition and their consequences. *The Journal of Politics*, 80(3), 964–981.
- Albertsen, A. (2022). How do the characteristics of Voting advice application users change over time? Evidence from the German Election Studies. *German Politics*, 31(3), 399–419.
- Benesch, C., Heim, R., Schelker, M., & Schmid, L. (2023). Do voting advice applications change political behavior? *The Journal of Politics*, 85(2), 684–700.
- Berwick, E., & Caughey, D. (2025). Modgirt: Multidimensional dynamic scaling of aggregate survey data. *Political Analysis*, 33(2), 91–106.
- Bornschier, S., Haffert, L., Häusermann, S., Steenbergen, M. R., & Zollinger, D. (2022, August 22–26). *Identity formation between structure and agency—How “us” and “them” relates to voting behavior in contexts of electoral realignment* [Conference paper]. ECPR General Conference, Innsbruck, Austria.
- Bornschier, S., Häusermann, S., Zollinger, D., & Colombo, C. (2021). How “us” and “them” relates to voting behavior—Social structure, social identities, and electoral choice. *Comparative Political Studies*, 54(12), 2087–2122.
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2020). *Cross-country trends in affective polarization* (Working paper No. 26669). National Bureau of Economic Research. <http://www.nber.org/papers/w26669>
- Caughey, D., O’Grady, T., & Warshaw, C. (2019). Policy ideology in European mass publics, 1981–2016. *American Political Science Review*, 113(3), 674–693.
- Cavaillé, C., & Trump, K.-S. (2015). The two facets of social policy preferences. *The Journal of Politics*, 77(1), 146–160.
- Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). Affective polarization, local contexts and public opinion in America. *Nature Human Behaviour*, 5(1), 28–38.
- Enyedi, Z. (2016). The influence of voting advice applications on preferences, loyalties and turnout: An experimental study. *Political Studies*, 64(4), 1000–1015.
- Fiorina, M. P., Abrams, S. A., & Pope, J. C. (2008). Polarization in the American public: Misconceptions and misreadings. *The Journal of Politics*, 70(2), 556–560.
- Fivaz, J., & Nadig, G. (2010). Impact of voting advice applications (VAAs) on voter turnout and their potential use for civic education. *Policy & Internet*, 2(4), 167–200.
- Frese, J., Hix, S., & Lachat, R. (2024). *Quality not quantity: How a VAA affected voting behavior in three large-scale field experiments*. OSF. <https://doi.org/10.31219/osf.io/y549s>
- Garzia, D., & Marschall, S. (2014). Voting advice applications in a comparative perspective: An introduction. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting Advice Applications in a comparative perspective* (pp. 93–104). Edward Elgar Publishing.
- Gemenis, K. (2024). Artificial intelligence and voting advice applications. *Frontiers in Political Science*, 6, Article 1286893.
- Gemenis, K., & Rosema, M. (2014). Voting advice applications and electoral turnout. *Electoral Studies*, 36, 281–289.
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis and interpretation*. WW Norton.
- Germann, M., & Gemenis, K. (2019). Getting out the vote with voting advice applications. *Political Communication*, 36(1), 149–170.

- Germann, M., Mendez, F., & Gemenis, K. (2023). Do voting advice applications affect party preferences? Evidence from field experiments in five European countries. *Political Communication*, 40(5), 596–614.
- Gidron, N., Adams, J., & Horne, W. (2019). Toward a comparative research agenda on affective polarization in mass publics. *APSA Comparative Politics Newsletter*, 29, 30–36.
- Gidron, N., Adams, J., & Horne, W. (2020). *American affective polarization in comparative perspective*. Cambridge University Press.
- Hajnal, A. (2025). Two tales of polarization? How affective and ideological polarization influence political corruption—A panel analysis of 153 countries between 2000 and 2021. *Political Studies*. Advance online publication. <https://doi.org/10.1177/00323217251324306>
- Heinsohn, T., Israel, J., Marschall, S., & Schultze, M. (2016). Online-Wahlhilfen in Wahlkämpfen. *Zeitschrift für Politikwissenschaft*, 26(3), 253–277.
- Hooghe, M., & Teepe, W. (2007). Party profiles on the web: An analysis of the logfiles of non-partisan interactive political internet sites in the 2003 and 2004 election campaigns in Belgium. *New Media & Society*, 9(6), 965–985.
- Iyengar, S., & Krupenkin, M. (2018). The strengthening of partisan affect. *Political Psychology*, 39, 201–218.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22(1), 129–146.
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405–431.
- Kingzette, J., Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). How affective polarization undermines support for democratic norms. *Public Opinion Quarterly*, 85(2), 663–677.
- Kleinnijenhuis, J., Van de Pol, J., van Hoof, A. M., & Krouwel, A. P. (2019). Genuine effects of vote advice applications on party choice: Filtering out factors that affect both the advice obtained and the vote. *Party Politics*, 25(3), 291–302.
- Ladner, A., Fivaz, J., & Pianzola, J. (2012). Voting advice applications and party choice: Evidence from SmartVote users in Switzerland. *International Journal of Electronic Governance*, 5(3/4), 367–387.
- 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.
- Leimgruber, P., Hangartner, D., & Leemann, L. (2010). Comparing candidates and citizens in the ideological space. *Swiss Political Science Review*, 16(3), 499–531.
- Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, 80(1), 392–410.
- Lelkes, Y. (2018). Affective polarization and ideological sorting: A reciprocal, albeit weak, relationship. *The Forum*, 16, 67–79.
- Levendusky, M. (2009). *The partisan sort: How liberals became democrats and conservatives became republicans*. University of Chicago Press.
- Levendusky, M., & Malhotra, N. (2016). Does media coverage of partisan polarization affect political attitudes? *Political Communication*, 33(2), 283–301.
- Levendusky, M. S. (2018a). Americans, not partisans: Can priming American national identity reduce affective polarization? *The Journal of Politics*, 80(1), 59–70.
- Levendusky, M. S. (2018b). When efforts to depolarize the electorate fail. *Public Opinion Quarterly*, 82(3), 583–592.
- Louwerse, T., & Rosema, M. (2014). The design effects of voting advice applications: Comparing methods of calculating matches. *Acta Politica*, 49(3), 286–312.

- Mahéo, V.-A. (2017). Information campaigns and (under) privileged citizens: An experiment on the differential effects of a voting advice application. *Political Communication*, 34(4), 511–529.
- Manavopoulos, V., Triga, V., Marschall, S., & Wurthmann, L. C. (2018). The impact of VAAs on (non-voting) aspects of political participation: Insights from panel data collected during the 2017 German federal elections campaign. *Statistics, Politics and Policy*, 9(2), 105–134.
- Marschall, S. (2014). Profiling users. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting Advice Applications in a comparative perspective* (pp. 93–104). Edward Elgar Publishing.
- Marschall, S., & Schultze, M. (2012). Voting advice applications and their effect on voter turnout: The case of the German Wahl-o-Mat. *International Journal of Electronic Governance*, 5(34), 349–366.
- Marschall, S., & Schultze, M. (2015). German e-campaigning and the emergence of a ‘digital voter’? An analysis of the users of the Wahl-o-Mat. *German politics*, 24(4), 525–541.
- Masket, S. E., & Noel, H. (2012). Serving two masters: Using referenda to assess partisan versus dyadic legislative representation. *Political Research Quarterly*, 65(1), 104–123.
- Mason, L. (2015). “i disrespectfully agree”: The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science*, 59(1), 128–145.
- Mason, L. (2018). Ideologues without issues: The polarizing consequences of ideological identities. *Public Opinion Quarterly*, 82(1), 866–887.
- McCarty, N. (2019). *Polarization: What everyone needs to know*. Oxford University Press.
- Michelitch, K., & Utych, S. (2018). Electoral cycle fluctuations in partisanship: Global evidence from eighty-six countries. *The Journal of Politics*, 80(2), 412–427.
- Munzert, S., Barbera, P., Guess, A., & Yang, J. (2020). Do online voter guides empower citizens? Evidence from a field experiment with digital trace data. *Public Opinion Quarterly*, 84(3), 675–698.
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706.
- Otjes, S., & Louwerse, T. (2014). Spatial models in voting advice applications. *Electoral Studies*, 36, 263–271.
- Pianzola, J. (2014). Selection biases in voting advice application research. *Electoral Studies*, 36, 272–280.
- Pianzola, J., Trechsel, A. H., Vassil, K., Schwerdt, G., & Alvarez, R. M. (2019). The impact of personalized information on vote intention: Evidence from a randomized field experiment. *The Journal of Politics*, 81(3), 833–847.
- Poole, K. T., & Rosenthal, H. (2000). *Congress: A political-economic history of roll call voting*. Oxford University Press.
- Poole, K. T., & Rosenthal, H. (2001). D-Nominate after 10 years: A comparative update to Congress: A political-economic history of roll-call voting. *Legislative Studies Quarterly*, 26, 5–29.
- Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, 16(1), 101–127.
- Riera, P., & Madariaga, A. G. (2023). Overlapping polarization: On the contextual determinants of the interplay between ideological and affective polarization. *Electoral Studies*, 84, Article 102628.
- Rogowski, J. C., & Sutherland, J. L. (2016). How ideology fuels affective polarization. *Political behavior*, 38(2), 485–508.
- Romero Moreno, G., Padilla, J., & Chueca, E. (2022). Learning VAA: A new method for matching users to parties in voting advice applications. *Journal of Elections, Public Opinion and Parties*, 32(2), 339–357.
- Rosema, M., Anderson, J., & Walgrave, S. (2014). The design, purpose, and effects of voting advice applications. *Electoral Studies*, 36, 240–243.
- Rosema, M., & Louwerse, T. (2016). Response scales in voting advice applications: Do different designs produce different outcomes? *Policy & Internet*, 8(4), 431–456.

- Schultze, M. (2014). Effects of voting advice applications (VAAs) on political knowledge about party positions. *Policy & Internet*, 6(1), 46–68.
- Sciarini, P. (2023). *Politique suisse: Institutions, acteurs, processus*. Presses Polytechniques et Universitaires Romandes.
- Selects. (2025). Swiss Election Study (Selects), cumulative dataset 1971-2023 (Version 3.0) [Data set]. FORS. <https://doi.org/10.48573/2w7s-1591>
- Sood, G., & Iyengar, S. (2016). *Coming to dislike your opponents: The polarizing impact of political campaigns* (Working paper). SSRN. <http://dx.doi.org/10.2139/ssrn.2840225>
- Stockinger, E., Maas, J., Talvitie, C., & Dignum, V. (2024). Trustworthiness of voting advice applications in Europe. *Ethics and Information Technology*, 26(3), Article 55.
- Stonecash, M. D. B., Jeffrey, M., & Mariani, M. D. (2003). *Diverging parties: Social change, realignment and party polarization*. Routledge.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup* (pp. 33–37). Brooks/Cole.
- Traber, D., Stoetzer, L. F., & Burri, T. (2023). Group-based public opinion polarisation in multi-party systems. *West European Politics*, 46(4), 652–677.
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15, 582–599.
- Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral studies*, 69, Article 102199.
- Walder, M. (2024). On the same space? Measuring the ideology of voters and political elites. *Frontiers in Political Science*, 6, Article 1425406.
- Walder, M. (2025). Latent ideological positions of Swiss parties and subsets of the population. *Swiss Political Science Review*. Advance online publication. <https://doi.org/10.1111/spsr.70003>
- Walder, M., Fivaz, J., Schwarz, D., & Giger, N. (2024). Explaining the (non-)use of voting advice applications. *Digital Government: Research and Practice*, 5(3), 1–29.
- Walgrave, S., Nuytemans, M., & Pepermans, K. (2009). Voting aid applications and the effect of statement selection. *West European Politics*, 32(6), 1161–1180.
- Walgrave, S., Van Aelst, P., & Nuytemans, M. (2008). 'Do the Vote Test': The electoral effects of a popular vote advice application at the 2004 Belgian elections. *Acta Politica*, 43(1), 50–70.
- Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the U.S. electorate. *American Politics Research*, 45(4), 621–647.
- Zollinger, D. (2024). Cleavage identities in voters' own words: Harnessing open-ended survey responses. *American Journal of Political Science*, 68(1), 139–159.

About the Authors



Maxime Walder (PhD) is a senior researcher at the University of Geneva and Basel. His research focuses on the behavioral aspects of politics, with a specific focus on political representation, ideology, and the comparison between political elites and citizens.



Jan Fivaz is a senior researcher at the Centre for Public Management at the University of Bern. He is a co-founder of the voting advice application Smartvote and has worked on research projects focusing on digital democracy, democratic innovations, and the impact of digitization on electoral decision-making, representation, and polarization.



Daniel Schwarz (PhD) is a senior researcher at the Centre for Public Management at the University of Bern. He is the co-founder of the voting advice application Smartvote. His research focuses on legislative behavior, political parties, digital democracy, and democratic innovations.

How Many People Use Voting Advice Applications? Survey Evidence From 15 Democracies Worldwide

Diego Garzia^{1,2}  and Frederico Ferreira da Silva³ 

¹ Institute of Political Studies, University of Lausanne, Switzerland

² Department of Political and Social Sciences, Bologna University, Italy

³ Institute of Social Sciences, University of Lisbon, Portugal

Correspondence: Diego Garzia (diego.garzia@unil.ch)

Submitted: 14 September 2025 **Accepted:** 26 November 2025 **Published:** 14 January 2026

Issue: This short note is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

In this short note, we offer an update on previous assessments of voting advice application (VAA) usage with representative samples from National Election Studies in established democracies worldwide. We map the spread of VAAs among the voting population in those countries in which post-election surveys asked their respondents whether they had used VAAs during the national election campaign. Our empirical base encompasses data from 48 elections held in 15 countries (Australia, Austria, Belgium, Canada, Denmark, Finland, Germany, Greece, Iceland, the Netherlands, New Zealand, Norway, Portugal, Sweden, and Switzerland) between 2003 and 2024. Beyond mapping VAA usage over time, our analysis focuses on the (potentially changing) individual-level correlates of VAA usage in a longitudinal perspective. The findings show a continuous increase in VAA usage figures in a large majority of the countries under analysis, but also provide little evidence of changes in the sociodemographic composition of the population of VAA users.

Keywords

diffusion theory; national election studies; survey data; voting advice applications

1. Introduction

The existence of voting advice applications (VAAs) as a research object is closely associated with their levels of diffusion among the electorates of contemporary liberal democracies. Arguably, the main reason to study VAAs and their effect is that they are widely used by citizens (Garzia, 2010; Garzia & Marschall, 2016; Marschall & Garzia, 2014). The Dutch *Stemwijzer* went from only 6,500 users in its first online edition of

1998 to over 9 million in 2023. The German Wahl-O-Mat has been available since 2002 and it has been used over 130 million times since then. In the campaign of the 2025 German Federal election alone it was used around 26 million times. Cross-national VAAs like EU&I are routinely used by millions of users during the European Parliament election campaign (Ferreira da Silva & Garzia, 2023; Reiljan et al., 2020). The growing popularity of these tools motivates increased research interest on their implications for political attitudes and behaviour.

However, usage numbers do not reflect the actual number of users. Citizens fill out either multiple VAAs during the same campaign (Dieing, 2025) or the same VAA's questionnaire multiple times, therefore inflating overall usage figures and VAA log files. Data cleaning techniques can mitigate this issue (Andreadis, 2014), but only to a certain extent. Survey data from national representative samples could, in turn, offer more accurate figures on the number of unique individuals using VAAs in each country. As VAAs gained relevance among the electorate, national election studies worldwide came to include questions about VAAs in their respective surveys. As yet another source of political information, they are frequently featured alongside other old and new media platforms to capture voters' levels of exposure to these online tools.

Initial evidence on the spread of VAAs among the electorate based on 21 national elections showed high levels of penetration among the electorates of Western European democracies, even if with important cross-country variation (Garzia & Marschall, 2019). In this short note, we expand this effort to map the proportion of VAA users in the voting-age population across a total of 48 elections from 15 democracies worldwide. The first goal is, thus, to provide an updated account of whether and to what extent VAAs have become increasingly used in these countries.

Additionally, we explore whether changes in usage numbers were accompanied by changes in the socio-political composition of VAA users. Previous research has traced a clear VAA-user profile: young, male, highly educated, strongly interested in politics, and predominantly left-wing oriented (Marschall, 2014; Walder et al., 2024). As VAAs become massified among the public, does the profile of VAA users come to resemble the characteristics of average population? Examining the changing profile of VAA users in Germany, Albertsen (2022, p. 399) finds a "development in which users become more similar to the population as a whole." The author develops a theoretical framework drawing on Rogers' (1962) innovation diffusion theory, according to which technological diffusion typically follows an early adoption by individuals differentiating in age and educational levels from mass-adopters, and finds evidence of a greater resemblance between VAA users and the average voter over time. Leveraging on the broad longitudinal and cross-national characteristics of our data, we thus propose an examination of the (changing) profile of VAA users in a large-N comparative perspective.

2. Data

Our analyses rely on post-election survey data from a total of 48 elections from 15 liberal democracies on three continents. These correspond to the overall amount of election studies worldwide for which we could identify variables on self-reported use of VAAs in the electoral campaign.

There is considerable variation across countries in the length of the time series. That is partially related to the differentiated timing of the introduction of VAAs in the different countries, but not exclusively. For example,

VAA questions had been customarily featured in the post-electoral waves of the German Longitudinal Election Study, but they were dropped after 2013.

It is important to note that, while we are convinced that national election studies remain the best data source on actual numbers of unique VAA users, this choice is still not without caveats. Just as we argued that research interest in VAAs is tied to their usage numbers, national election studies are also more prone to include questions about VAAs only once they permeate a significant portion of the electorate. Hence, since a necessary pre-condition for inclusion is a certain degree of popularity, our sample becomes less likely to include cases in which VAAs are (still) used in lower numbers. Nevertheless, the absence of VAA questions does not necessarily signal a demand-side issue. Italy, for example, is consistently among the top-scoring countries in pan-European VAAs for the European Parliament elections (Michel et al., 2019). Since popular VAAs are virtually nonexistent at the national level, they are nonetheless absent from the respective election studies.

Question wording naturally varies substantially across countries and over time (see Table A1 of the Supplementary File). In some countries, VAAs are referred to using generic terms such as vote compass, vote matcher, voting advice website, election quiz, or candidate selectors. In other countries, they are referred to by the actual name of the VAA (e.g., stemtest in Belgium, smartvote in Switzerland, and Wahl-O-Mat in Germany). This also speaks to the relative popularity and diversity of these tools across countries.

3. Results

Figure 1 presents self-reported levels of VAA usage in the respective country samples. The data shows wide usage of VAAs in contemporary elections among the publics of the 15 countries analysed. The figures for the most recent elections available (post-2020) range from 20% in the 2023 Swiss election to 68% in Finland in the same year. That is nearly as many people using a VAA as voting in the 2023 Finnish election (official turnout rate was 72.6%).

Not only are VAAs widely used, but the number of users has also been increasing in virtually every country analysed. Finland has registered a nearly threefold increase in VAA users over the last two decades; the Netherlands has doubled the number of VAA users across a similar time period; Denmark, Norway, and Sweden have doubled them in little more than a decade. In Scandinavia and the Netherlands, VAAs are used by the majority of the voting age population in recent elections.

Naturally, VAAs tend to be more used in the countries where they have been implemented for longer. However, countries like Germany or Switzerland have had reputable VAAs for nearly two decades now, and the latest figures available are on the lower end. The size of the party system, on the other hand, seems to correlate strongly with aggregate usage numbers. In line with what we know from previous comparative research on VAAs, their usage increases with the complexity of the voting task stemming from larger party systems. Other party system features, such as the existence of candidate-based elections, do not seem to go hand in hand with high VAA usage numbers. In fact, the smaller multi-party systems of German-speaking countries, despite being among the pioneers, did not increase over time as fast as Anglo-Saxon countries, like Australia and New Zealand, which, despite introducing these tools later, now report usage figures more or less in line with Austria, Germany, or Switzerland. Finally, attempts at measuring VAAs in South European

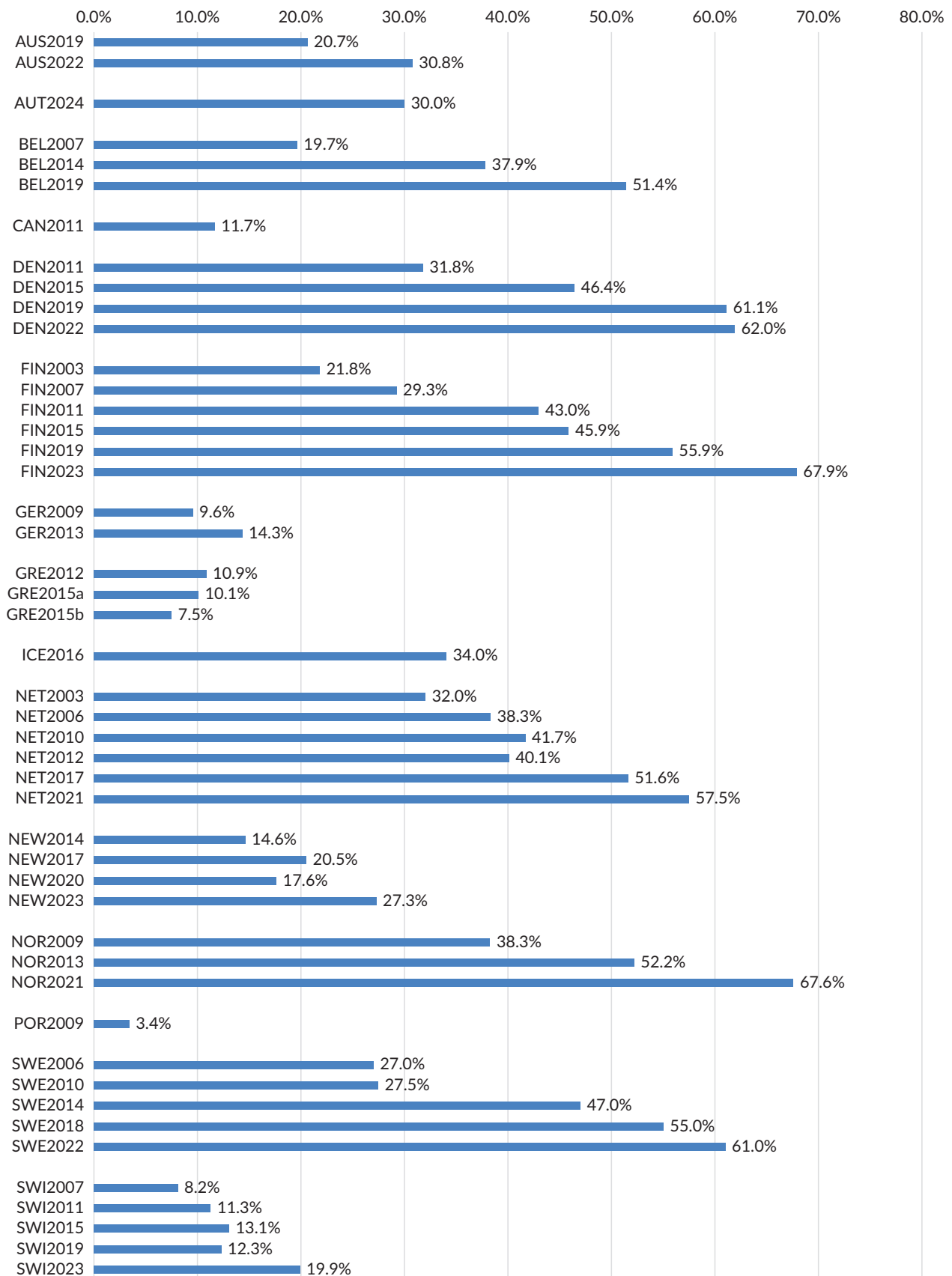


Figure 1. Percentage of self-declared VAA users in national election study datasets.

countries, like Greece or Portugal, have not been repeated in the most recent election studies, leaving us with an incomplete picture for that region.

To examine whether the young, male, left-leaning, highly educated, and interested in politics profile of VAA users identified in the literature (Marschall, 2014) still holds, we estimated logistic regression models of VAA usage (full Model 1 estimates are reported in Table A2 of the Supplementary File). We began by pooling all the valid observations at our disposal, considering all countries and elections. The dependent variable is *self-reported VAA usage during the election campaign* (0. No; 1. Yes). Independent variables are *age* (numeric), *gender* (0. Male; 1. Female), *education* (1. Primary; 2. Secondary; 3. Tertiary), *interest in politics* (1. Not interested, 2. Neither interested nor uninterested, 3. Interested), *left-right self-placement* (0. Left; 10. Right), and *strength of party identification* (0. None, 1. Weak, 2. Fairly strong, 3. Very strong). The model also includes country-fixed effects and controls for the year in which the survey was fielded.

The results largely confirm previous findings from the literature on users' profiling. Figure 2 presents the average marginal effects (AMEs) for all independent variables. According to the model, age is strongly negatively associated with the probability of using a VAA: a 60-year-old is about 20 percentage points less likely to do so than a 30-year-old. Gender differences are small and not statistically significant, while education and political interest show marked positive effects. Each step up in education increases the likelihood of VAA use by roughly 10 percentage points (around 20 points between primary and tertiary education), while moving from low to high political interest adds about 15 points. By contrast, ideological position and partisan strength have negative effects. Individuals located at the extreme left of the ideological scale are about 4 points more likely to use a VAA than those at the extreme right, and respondents with very strong party identification are roughly 2 percentage points less likely than those without any partisan attachment.

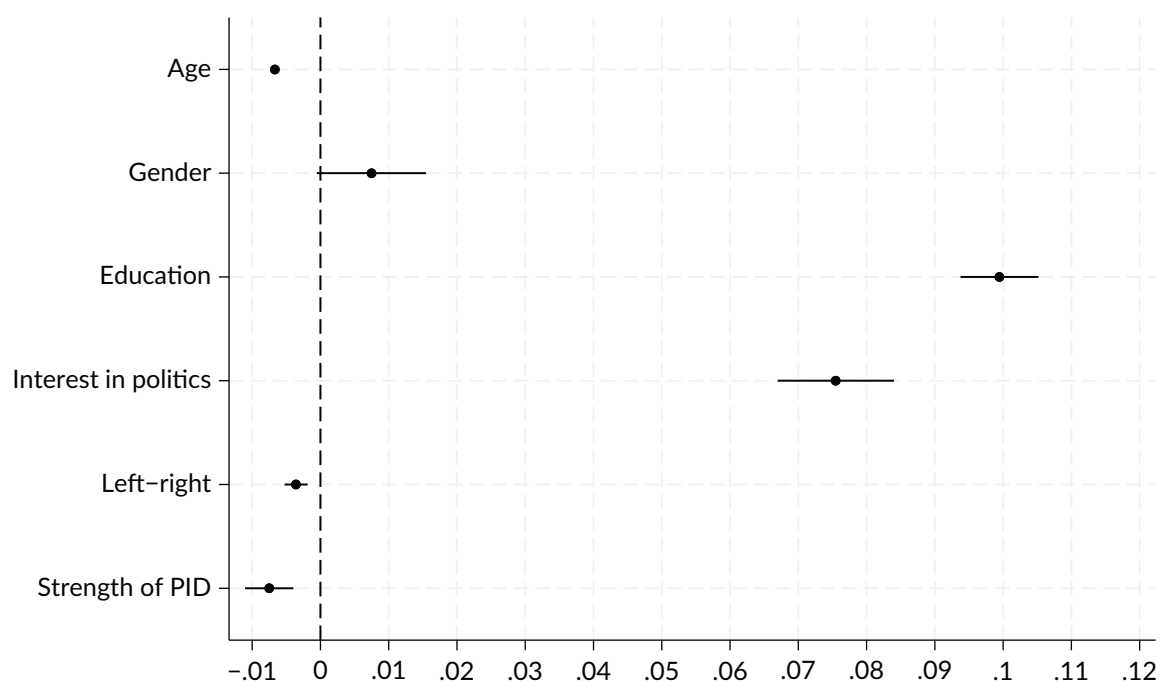


Figure 2. Individual-level correlates of VAA usage: AMEs from a logistic regression model with country fixed effects.

We have then interacted each predictor with the election year to capture potential variations in their predictive ability over time. In the naïve interaction model (full Model 2 estimates are reported in Table A2 of the Supplementary File), we find that, over time, age, education, and interest in politics became less relevant in explaining VAA usage. A leave-one-out cross-validation analysis of these results reveals that, however, the pooled model is sensitive to the sample composition (see Figure A1 in the Supplementary File). Only two interaction effects remain robust to our sensitivity tests. As a result, in Figures 3 and 4, we present the changing AMEs of the two interactions that are consistently robust to model specification, namely, age and interest in politics.

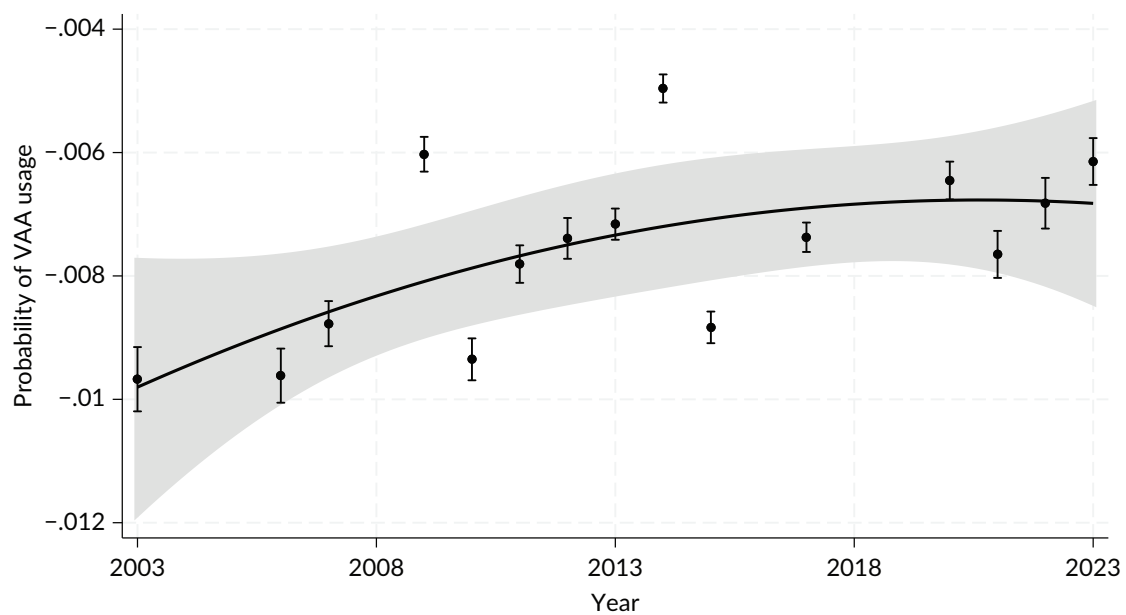


Figure 3. The changing relationship between age and VAA usage over time: Interaction effects.

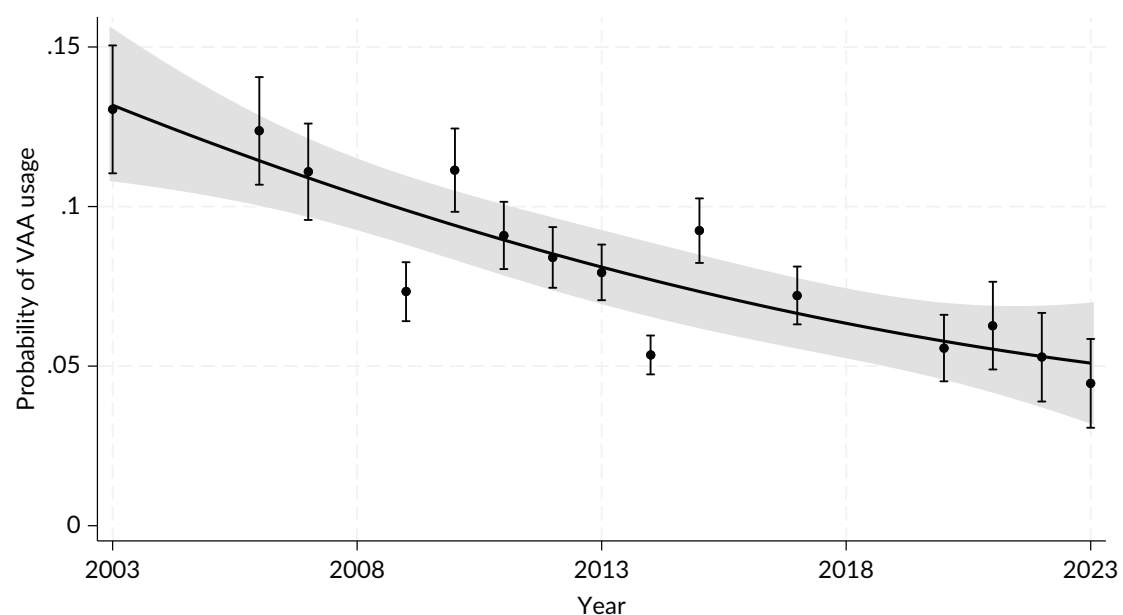


Figure 4. The changing relationship between interest in politics and VAA usage over time: Interaction effects.

To test the robustness of our findings, we have also estimated a multilevel model including the effective number of electoral parties for each election and a variable contrasting proportional representation with first-past-the-post electoral systems to account for potential structural differences in patterns of VAA usage linked with institutional features of the voting system (see Table A3 of the Supplementary File). Their inclusion does not substantively alter the findings reported in the main text.

4. Conclusion

The analysis has demonstrated that VAAs are heavily and increasingly used in liberal democratic multiparty systems worldwide. However, their spread among the electorate does not correspond to a substantial change in the sociopolitical profile of their users. We do witness a decreasing predictive power of age and interest in politics as correlates of VAA usage across the last two decades. However, and most importantly, we find that in spite of their massive spread across the electorate, VAA users remain predominantly young, highly educated, interested in politics, and left-leaning. These findings carry clear implications for future research on the dynamics of VAA effects on political behaviour in comparative and longitudinal perspectives.

Acknowledgments

We are grateful to Micha Germann, Henrik Oscarsson, Elie Michel, and Michiel Nuytemans for sharing information on VAA usage figures from election studies currently under embargo.

Funding

This work was supported by the Swiss National Science Foundation [grant n. PCEFP1_186898 to D.G.]. Publication of this short note in open access was made possible through the institutional membership agreement between the University of Lausanne and Cogitatio Press.

Conflict of Interests

In this short note, editorial decisions were undertaken by Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University).

Data Availability

The data that support the findings of this study are openly available. Our dataset is composed of 45 merged National Election Study, all available online with the national producing institutions.

Supplementary Material

Supplementary material for this short note is available online in the format provided by the authors (unedited).

References

- Albertsen, A. (2022). How do the characteristics of voting advice application users change over time? Evidence from the German election studies. *German Politics*, 31(3), 399–419.
- Andreadis, I. (2014). Data quality and data cleaning. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting advice applications in comparative perspective* (pp. 79–91). ECPR Press.
- Dieing, T. I. (2025). Just can't get enough—Profiling users of multiple voting advice applications. *Journal of Information Technology & Politics*. Advance online publication. <https://doi.org/10.1080/19331681.2025.2496259>

- Ferreira da Silva, F., & Garzia, D. (2023). Voting advice applications. In N. Carter, D. Keith, G. Sindre & S. Vasilopoulou (Eds.), *The Routledge handbook of political parties* (pp. 450–465). Routledge.
- Garzia, D. (2010). The effects of VAAs on users' voting behaviour: An overview. In L. Cedroni & D. Garzia (Eds.), *Voting advice applications in Europe. The state of the art* (pp. 13–33). Civis.
- Garzia, D., & Marschall, S. (2016). Research on voting advice applications: State of the art and future directions. *Policy & Internet*, 8(4), 376–390.
- Garzia, D., & Marschall, S. (2019). Voting advice applications. In W. Thompson (Ed.), *Oxford research encyclopedia of politics*. Oxford University Press.
- Marschall, S. (2014). Profiling users. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting advice applications in comparative perspective* (pp. 93–104). ECPR Press.
- Marschall, S., & Garzia, D. (2014). Voting advice applications in a comparative perspective. An introduction. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting advice applications in comparative perspective* (pp. 1–10). ECPR Press.
- Michel, E., Cicchi, L., Garzia, D., Ferreira da Silva, F., & Trechsel, A. H. (2019). *euandi2019: Project description and datasets documentation* (Research paper no. RSCAS 2019/61). European University Institute.
- Reiljan, A., Ferreira da Silva, F., Cicchi, L., Garzia, D., & Trechsel, A. H. (2020). Longitudinal dataset of political issue-positions of 411 parties across 28 European countries (2009–2019) from voting advice applications EU Profiler and euandi. *Data in Brief*, 31, Article 105968.
- Rogers, E. M. (1962). *Diffusion of Innovations*. Free Press.
- Walder, M., Fivaz, J., Schwarz, D., & Giger, N. (2024). Explaining the (non-) use of voting advice applications. *Digital Government: Research and Practice*, 5(3), 1–29.

About the Authors

Diego Garzia is an assistant professor of political science at the University of Lausanne, Switzerland. His research interests focus on elections, public opinion, and voting behavior in a comparative perspective. He currently serves on the steering committee of the Consortium for National Election Studies.

Frederico Ferreira da Silva is a research fellow at the University of Lisbon, Portugal. He works on elections, public opinion, and voting behavior in Western democracies, with a recent interest in affective polarization in a comparative and longitudinal perspective.

Exploring Their Options? Tracking How Voters Actually Use Voting Advice Applications

Andreas Videbæk Jensen ¹ , Veikko Isotalo ² , and Mathias Wessel Tromborg ³ 

¹ Department of Political Science, Aarhus University, Denmark

² Faculty of Social Sciences, University of Helsinki, Finland

³ Department of Political Science, Aarhus University, Denmark

Correspondence: Mathias Wessel Tromborg (tromborg@ps.au.dk)

Submitted: 2 September 2025 **Accepted:** 18 November 2025 **Published:** 14 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Voting advice applications (VAAs) are a widely used tool by voters across democracies. However, we know surprisingly little about how voters actually interact with these tools in practice. Here, we track the objective usage behavior of a consenting and representative sample of voters in an online VAA in Denmark. This enables us to identify the extent to which VAA users explore their positional congruence with more than one of their electoral options as well as how much time they spend doing so. Using this data, we find that a majority of users focus on their own party, but also that a substantial number of users explore other party options and spend significant time doing so. Moreover, these patterns are conditioned by vote choice certainty and political interest. Undecided and uncertain voters are more likely to explore multiple party options, and politically interested users engage with these features as well. The results have important implications for research on VAA effects, which has so far predominantly focused on VAA advice but not on the potential impact of other VAA functionalities. Moreover, the implication for practitioners is that there is a demand for VAA functions that allow users to obtain more nuanced information than merely the voting advice.

Keywords

behavioral data; elections; political behavior; vote choice; voting; voting advice applications

1. Introduction

Voting advice applications (VAAs) are popular tools in contemporary democracies with millions of users in countries such as the Netherlands, Finland, Germany, and Denmark (Bundeszentrale für politische Bildung,

2021; Hansen & Stubager, 2019; Statistics Finland, 2023; van de Pol et al., 2019). It is thus not surprising that VAAs have attracted scholarly attention among researchers. So far, much of the research has focused on whether and how VAAs influence their users' vote intention and choice. Most experimental studies find no effects of VAAs use on vote switching (Enyedi, 2016; Mahéo, 2016; Munzert et al., 2020; Pianzola et al., 2019), but more recent research has shown that this is likely because VAAs effects depend on the type of advice that is given (Germann et al., 2023; Tromborg & Albertsen, 2023). Some users receive voting advice that is congruent with their pre-existing party preference, which strengthens their intention to vote for this party, while others receive incongruent advice, which leads them to switch to the recommended party instead. These effects are especially strong among voters who are politically interested enough to use a VAA but not so interested that they engage in more time-consuming efforts, such as following political news and debates. Undecided voters are also more likely to follow the advice (Garry et al., 2019; Germann et al., 2023; Tromborg & Albertsen, 2023).

While we know a great deal about the effects of voting advice, we still know surprisingly little about how users engage with VAAs. Existing VAAs research on vote intention and choice has focused almost exclusively on the effects of the VAA recommendation, but users may also seek and use information from other VAAs features. For example, it has been argued that more politically interested users are less influenced by VAAs advice because they seek more nuanced information elsewhere (Tromborg & Albertsen, 2023). Instead, they use VAAs to confirm their choice or for entertainment purposes (van de Pol et al., 2014). However, VAAs can also provide relatively nuanced information—for example, by showing users party positions on each VAA question along with party statements about why they take the issue positions that they do. Some politically interested users may engage with this information because it provides information that they find relevant. Likewise, voters who use VAAs to decide whom to vote for may engage with this additional information across several of their party options to receive potentially relevant information. We refer to citizens who access this type of information as “explorers.”

The purpose of this article is to understand how much voters engage with nuanced VAA functionalities and to map which types of voters are more likely to use these functionalities. To do so, we collected a consenting and demographically representative sample of Danish voters, asked them to use a VAA, and answer a set of background questions. We recorded how voters interacted with the VAA in terms of which buttons they clicked, what parties they explored, and how much time they spent on various activities in the VAA. This allows us to map the extent to which voters explore party positions on individual issues and the party explanations for taking those positions. Moreover, we measure the survey respondents' political interest, vote choice certainty, and self-reported motivation for using VAAs. This allows us to analyze how different types of people use VAAs differently, as well as their reasons for doing so.

We find that most of our respondents did not engage with more nuanced VAAs information than they had to by design, but still, a substantial share did. A little more than a quarter (28%) of the respondents sought out the detailed issue stances of more than one party, even though they did not have to, and an additional 12% read the parties' own statements about why they took the positions that they did. These numbers are likely conservative estimates of real-life VAAs behavior, since the respondents in our sample are paid to take surveys and they received no monetary benefit for exploring these options. Moreover, we find that undecided and uncertain voters are more likely to report using VAAs to find out whom to vote for and to explore the issue positions of more than one party. Furthermore, politically interested VAAs users, if anything, are more likely to explore multiple party options than those who are less politically interested.

These results have important implications for VAAs researchers and practitioners. In terms of research, they corroborate previous findings that undecided voters are more likely to be influenced by a VAA. Furthermore, they demonstrate that the information role of VAAs goes well beyond an advisory role. Many users engage with the more nuanced information that is frequently available in VAAs after the advice has been given, and especially so when they are undecided or uncertain about whom to vote for. Moreover, politically interested voters do not just use VAAs to confirm their choice—many of them also use VAAs to explore other options. This indicates that existing research on VAAs effects likely underestimates the total influence that these tools have on voters. A key implication for VAAs designers is that there is good reason to include user options to explore detailed information because a non-trivial number of users engage with these additional information sources.

2. Background

Much of our knowledge on how voters use VAAs comes from van de Pol et al.'s (2014) study on Dutch VAAs users during the 2012 general election in the Netherlands. In this study, 7% of the users of the Kieskompas VAA opted in to take a post-VAA survey. The authors used a latent class analysis on their behavior to identify three types of VAA users: doubters, seekers, and checkers. Doubters and seekers both use VAAs to find out whom to vote for, but differ in their level of certainty about which party to vote for prior to using the VAA (with seekers having clearer party preferences). Checkers, on the other hand, have already decided and use the VAA primarily to confirm their choice or for entertainment purposes. This group was the largest of the three ($\approx 58\%$ of the sample) and was characterized by a high level of political interest. More recently, it has been suggested by Dieing (2025) that there is an additional group of users called skeptics, who are more prone to use only a single VAA instead of multiple and are supportive of right-leaning ideology, showing dissatisfaction with the political system.

While this typology has contributed greatly to our understanding of how different types of citizens intend to use VAAs, our study contributes in several ways. First, we analyze how voters use VAAs in a different country and at a more recent time. This is important to achieve a more general understanding of VAAs users in a more contemporary context. Second, we use a sample of demographically representative citizens. Thus, we can say how the average voter would use a VAA. This is a useful contribution because, as van de Pol et al. (2014) note, the Dutch citizens who opted in to the post-VAA survey were systematically different from other users and citizens. Third, we include objective behavioral measures of how users interacted with the VAA, including post-advice options. In conjunction with self-reported data on the users' political interest, vote certainty, and reasons for using a VAA, this allows us to assess whether VAAs users with high levels of interest and certainty are merely checkers, or whether they also use other VAA information features.

Based on the results from the recent behavioral VAA research reviewed in the introduction, we expect that citizens with low political interest mainly focus on the advice when using a VAA. These users need a simple, low-cost cue that enables them to make a more informed choice than they otherwise would (Tromborg & Albertsen, 2023), and a recommendation from a VAA has the potential to provide exactly that. For undecided or uncertain users, we expect the picture to be different. These users are not necessarily uninterested in politics and need quality information that allows them to make the best choice given their preferences. Consequently, they should be more likely to engage with more time-consuming VAA information features than decided voters on average. Finally, we also expect that high-interest voters are not merely checkers but that they also engage more with nuanced VAA information. Existing research

suggests that the effects of VAA advice are negligible among citizens in this latter group, which makes sense considering that they can gain more nuanced information about their candidates and parties elsewhere (e.g., in the news, following debates, and social media). Likewise, van de Pol et al.'s (2014) results suggest that they primarily use VAAs to confirm their choice or for entertainment. However, if politically interested voters do, in fact, demand nuanced information, then they may perceive an incentive to engage with more detailed VAA information features, such as exploring party positions on individual issues. In this way, VAAs may not just influence low-interest voters as an information shortcut but also high-interest voters as a more detailed information tool.

3. Case, Data, and Methods

The analyses draw on data collected in Denmark, where VAAs are very popular, with around 60% of the electorate actively using one or more VAAs in a general election (Hansen & Stubager, 2019). In Denmark, VAAs are mostly designed and implemented by media outlets, such as newspapers and TV stations. In this study, we draw on a VAA developed by us in collaboration with a major Danish newspaper. The mentioned VAA used in the study is a party-based VAA that is running between elections on the newspaper's website (jp.dk). The VAA consisted of three parts (Figures B1–B7 in the Supplementary File depict the visual appearance of the VAA): First, users were asked to submit their answers on a set of issue positions—the so-called VAA statements. An example of such a statement is shown in Figure B4 in the Supplementary File. Note that the current VAA includes 18 issue statements. This is slightly fewer than most Danish VAAs, which means that we may underestimate how much time voters generally spend filling out the statements. Second, users proceeded to the results page of the VAA, where they were shown which party's issue stances were most congruent with their own positions. They were also shown the top five and bottom five of these matches (Figure B5 in the Supplementary File displays an example of the results page). The VAA includes 12 different parties listed in Section D of the Supplementary File. The third part of the VAA was the detailed party profiles. By clicking on one of the parties displayed on the results page, a detailed overview of this party's issue positions would be shown, including the option to read party statements about why they took the stance that they did. An example of a detailed party profile is shown in Figure B6 in the Supplementary File. Users could also view parties' written justifications for their policy positions (see Figure B7 in the Supplementary File). After viewing any party profile, users could then return to the results page and open another party profile if they wanted to. Importantly, in this study, we designed the VAA so that voters had to remain on the results page for at least 30 seconds, and they had to open at least one party profile to proceed. This means that when we look at the time users spent on results and party profiles—and when we look at the number of party profiles users opened—it is important to keep this baseline requirement in mind, as we highlight in Section 4. To assess the extent to which these requirements limit the generalizability of the findings in terms of what kind of parties voters choose to visit, we report both the findings for the entire sample and the sample of users who visit more than one party. For this latter group, we know that they would visit party profiles even if the requirements were not there.

3.1. Data Collection and Analyses

We recruited an online panel of respondents through the survey company Epinion and directed respondents to the VAA described above. Moreover, we asked them to fill out a background survey. The sample was selected to meet census-based quotas on gender, age, region, and education. The users were informed

about the study and gave active consent. Accordingly, we collected two kinds of data: First, we collected data on the respondents' background characteristics, political interest, vote certainty (in a pre-VAA survey), and self-reported motivation for using a VAA (in a post-VAA survey). To measure vote certainty, we asked respondents which party they would vote for if there were a parliamentary election tomorrow. *Don't know* responses were coded as undecided. If a voter selected a party, they were asked how certain their choice was (possible responses were *not at all certain*, *quite certain*, or *very certain*). To measure political interest, respondents were asked: "How interested in politics would you say that you are?" Response options were *not at all interested*, *not very interested*, *somewhat interested*, or *very interested*. Second, we recorded how the respondents interacted with the VAA in terms of the options they clicked, the parties they explored, and how much time they spent on the various activities in the VAA. This rich set of objective behavioral data constitutes a key part of this article's contribution and is summarized in Figure A.1 in the Supplementary File.

The collected survey data contain two groups of respondents: those who were asked to use a VAA in addition to a survey ($n = 1,011$) and those who only answered a survey ($n = 1,091$). We focus our analysis on the first group because this allows us to combine respondents' background characteristics and VAA attitudes to their actual VAA behavior. VAA behavior was tracked by recording how users responded to VAA statements and how they navigated through the VAA results after having received the advice.

It is important to note that our study is based on participants who would use a VAA at their own initiative and participants who would not. Specifically, 63% of participants reported that they had used a VAA before (cf. Table A.6 in the Supplementary File). On the one hand, our study thereby avoids the potential idiosyncratic nature of a sample based on a specific VAA from a specific provider with a specific audience. On the other hand, the fact that not all of the respondents for whom we track how they use the VAA would actually use a VAA in their own time means that our data captures a combination of actual VAA use and how voters who do not use VAAs *would* use a VAA. This is a limitation, but the latter component also has value in itself, since VAA usage is increasing (see Garzia & Ferreira da Silva, 2026). Similarly, the data on motivations for using a VAA also capture actual motivations as well as why some voters would use a VAA at their own initiative, even though they have not done so yet.

4. Results

We present three sets of results: First, we show descriptive statistics that reveal how our users generally interact with VAAs using self-reported and behavioral data. Second, we analyze how these general patterns are associated with political interest. Third, we replicate the analyses from the second section but substitute political interest with vote choice certainty.

4.1. How Do Voters Interact with VAAs in General?

In this section, we present descriptive statistics on how users interact with a VAA. First, we describe the reasons the respondents themselves gave for using VAAs. Second, we turn to our behavioral measures and illustrate how much time voters spent filling out the VAA statements and, subsequently, exploring different party profiles. Third, we report the number and types of parties that users were interested in viewing detailed information about.

4.1.1. Voters' Self-Reported Motivations for Using VAAs

Table 1 summarizes what voters reported as their main motivation for using a VAA. We use the response options from van de Pol et al. (2014) and compare the responses in this study to ours. In doing so, we observe a much smaller proportion of respondents reporting that they use VAAs for entertainment (i.e., responding that it is an interesting test to reflect and discuss) and an increase in wishing to learn about parties' opinions. It is unclear whether those differences are due to the country context, the timing, or the representativeness of the sample. Regardless, the result is interesting because these two responses characterize checkers in the framework developed by van de Pol et al. In other words, we observe much less "checking" behavior in our sample compared to the Dutch sample from 2014. Instead, a majority of our respondents report using VAAs to learn about parties' opinions or to find out whom to vote for.

Table 1. Users' self-reported motivations for using a VAA.

Sample	Learn about parties' opinions	Check if my party agrees with me	Find out who/ what to vote for	Interesting test to reflect and discuss	Obs.
van de Pol et al. (2014)	16%	39%	17%	28%	52,999
Danish 2015 sample	29% (26%–31%)	34% (31%–37%)	25% (22%–28%)	12% (10%–14%)	1,011

Note: Cells report proportions with 95% confidence intervals in parentheses.

4.1.2. How Much Time Do Users Spend on Various Parts of the VAA?

We now turn to the respondents' actual behavior while using the VAA. As mentioned in Section 3, the VAA consisted of three parts: First, voters answered issue statements in order for the VAA algorithm to calculate the congruence between their stances and those of the Danish parties. Next, users were presented with the results, showing which parties the users were most congruent with on the VAA issues. On this page, users can click on the different parties to see the detailed issue profile for the party. Time spent viewing parties was included in the time spent on the results page. The study was designed so that voters had to remain on the results page for at least 30 seconds, and they had to open at least one party profile to proceed. As reported in the Supplementary File (Figure A.2) the most time was spent on filling out the issue positions (18 statements in total). The median user spent a bit less than 2.5 minutes on this task (144 seconds). Second, we see that users spent significantly more time on the results page (which includes viewing party profiles) than they were required to. The median user spent just above 1.5 minutes (92 seconds), with 50% of all users spending between 55 and 150 seconds. This indicates that users are quite interested in knowing which parties they agree with the most, also beyond the highest match. Finally, more than half of the users spent more than the required amount of time viewing one or more detailed party profiles. The median time spent on viewing party profiles was 40 seconds, with 50% of voters spending between 16 and 88 seconds there.

4.1.3. How Many and Which Party Profiles Do Users Open?

In terms of the overall number of parties viewed, 72.5% of the sample viewed only one party, and 27.5% viewed two or more. When interpreting these numbers, it is important to remember that participants had to open one profile to proceed in the survey. It is likely that many among the 72.5% would not have visited any

party profiles. However, these numbers tell us that about one-fourth of the participants explored at least one more party than what was mandatory. This group did not have to view the second party that they viewed. Thus, we can assume that they were legitimately interested in viewing more than one party. In other words, one-fourth of the users were not simply looking for a recommendation or some details on the recommended party's positions, nor were they simply checkers that just wanted to know about the positions of the party they initially intended to vote for. Instead, they explored multiple party options. We also see that this group spent more time viewing party profiles. In fact, the median time spent viewing party details was 36 seconds for the users who viewed just one party, but 66 seconds for those who looked at more than one party. More specifically, the median time for those who viewed two parties was 58 seconds. For three and four parties, it was 87.5 seconds and 103 seconds, respectively. The time spent viewing party profiles was higher by between 15 and 30 seconds for each additional party viewed (Figure A.2 in the Supplementary File).

Next, we report which party options the users chose to read about after receiving their VAA recommendation. van de Pol et al. (2014) find that a majority of users are checkers who want to see if their intended party choice agrees with them. Consistent with this, we see in Table 2 that a majority of users open the issue profile of the party they intended to vote for before taking the VAA, but also that a substantial percentage of users do not (45%). This distribution is similar when we look at those who open only one party profile compared to those who explore several parties, with 53% and 60% consulting the profile of their pre-VAA party choice, respectively. While a majority of users thus check their own party, a substantial proportion of users do not, and explore one or more parties they did not expect to vote for. Note that for 19% of the sample, the intended party choice was the same as the party recommended by the VAA.

Table 2. Number of parties opened and whether the user opened the party profile of the party they intended to vote for before taking the VAA.

Group	Did not open pre-vote party	Opened pre-vote party	Total
Opened one party profile	47% (43%–50%)	53% (50%–57%)	100% (n = 733) [72.5% of all users]
Opened several party profiles	40% (34%–46%)	60% (54%–66%)	100% (n = 278) [27.5% of all users]
All users	45% (42%–48%)	55% (52%–58%)	100% (n = 1,011)

Notes: Cells report row percentages with 95% CIs in parentheses; Pearson's χ^2 test $p = 0.063$.

Finally, we turn to the feature where users can see the parties' detailed explanations for their issue stances. Specifically, users had the option to click a button to see the party's own written explanation of its position on each of the issue statements. This button was used by 12.1% of the users (Table 3). This means that seven out of eight users did not use this feature. However, if we compare those who only opened one party profile to those who opened several party profiles, the latter group was twice as likely to use the "see written explanations" feature (about one in five vs. one in 10). Clearly, this group exhibits "exploring"-behavior.

Table 3. Share of users opening the parties' own written explanations.

Group	Did not open written explanations	Opened written explanations	Total
Opened one party profile	90% (88%–93%)	10% (7%–12%)	100% (n = 733)
Opened several party profiles	81% (77%–86%)	19% (14%–23%)	100% (n = 278)
All users	88% (86%–90%)	12% (10%–14%)	100% (n = 1,011)

Notes: Cells report row percentages with 95% CIs in parentheses; Pearson's χ^2 test $p < 0.001$.

4.2. Which User Characteristics Predict How Voters Use VAAs?

In this section, we analyze whether and how political interest and vote choice certainty condition the ways voters use VAAs. We present these analyses in two parts: The first part examines how political interest conditioned the respondents' self-reported reasons for using VAAs as well as their actual behavior while using the VAA. The second part replicates this analysis but substitutes political interest with vote choice certainty.

4.2.1. Does Political Interest Predict How Users Interact With VAAs?

A key expectation from van de Pol et al. (2014) is that politically interested voters are more likely to use VAAs as checkers who compare their own views to those of their preferred party or for entertainment purposes. Less interested voters, on the other hand, are more likely to use VAAs to find out whom to vote for. When looking at users' self-reported motivations for using VAAs in Table 4, there is clear evidence for the latter but mixed evidence for the former. Less interested voters are clearly more likely to report using VAAs to find out whom to vote for. In fact, the least interested voters are more than three times as likely to use this reason compared to the most interested voters (46% and 14%, respectively). However, while highly interested users are more likely than other users to report that they use VAAs for entertainment, they are no more likely than

Table 4. Political interest and self-reported motivations for using VAAs.

Political interest	Learn about parties' opinions	Check if my party agrees with me	Find out who/what to vote for	Interesting test to reflect/discuss	Total
Not at all interested	21% (13%–29%)	23% (15%–32%)	46% (36%–57%)	10% (3%–15%)	100% (n = 95)
Not very interested	27% (22%–31%)	32% (27%–37%)	31% (26%–36%)	10% (7%–13%)	100% (n = 336)
Somewhat interested	31% (27%–36%)	38% (33%–42%)	19% (15%–23%)	12% (9%–15%)	100% (n = 416)
Very interested	31% (24%–38%)	35% (27%–42%)	14% (9%–20%)	20% (14%–26%)	100% (n = 161)
All users	29% (26%–31%)	34% (31%–37%)	25% (22%–28%)	12% (10%–14%)	100% (n = 1,008)

Note: Pearson's χ^2 test $p < 0.001$.

moderately interested users to say that they use VAAs to check if their party agrees with them. Likewise, they are no less likely to report that they use VAAs to learn about the parties' opinions. This indicates that while some politically interested users are certainly checkers, a substantial number of them are not.

Politically interested voters do, however, show substantially more checking behavior when we measure their actual engagement with the VAA (Table 5). In fact, 71% of very interested participants checked the profile of the party they expected to vote for prior to using the VAA, while only a minority of those with no or only little political interest did so (23% and 48%, respectively).

Table 5. Political interest and whether the user opened the party profile of the party they intended to vote for before taking the VAA.

How interested would you say you are in politics?	Did not open pre-vote party profile	Opened pre-vote party profile	Total
Not at all interested	77% (68%–85%)	23% (15%–32%)	100% (n = 95)
Not very interested	52% (46%–57%)	48% (43%–54%)	100% (n = 336)
Somewhat interested	38% (34%–43%)	62% (57%–66%)	100% (n = 416)
Very interested	29% (22%–36%)	71% (64%–78%)	100% (n = 161)
All users	45% (42%–48%)	55% (52%–58%)	100% (n = 1,008)

Note: Cells report row percentages with 95% CIs in parentheses; Pearson's χ^2 test $p < 0.001$.

However, checking behavior does not exclude exploring behavior. To assess whether politically interested voters, in addition to checking, also use VAAs for exploring, Table 6 reports the respondents' propensity to open one or more party profiles by their political interest. The results clearly demonstrate that politically interested users are not merely checkers. In fact, if anything, they are more likely to open several party profiles than less interested users, with 32% of the most interested users exploring multiple parties (as opposed to 21% among the least interested). However, the differences are not statistically significant. What we see is

Table 6. Political interest and opening one or several party profiles.

Political interest	Opened one party profile	Opened several party profiles	Total
Not at all interested	79% (71%–87%)	21% (13%–29%)	100% (n = 95)
Not very interested	71% (66.6%–76.3%)	29% (23.7%–33.4%)	100% (n = 336)
Somewhat interested	74% (69.3%–77.8%)	26% (22.2%–30.7%)	100% (n = 416)
Very interested	68% (60%–75%)	32% (25%–40%)	100% (n = 161)
All users	72% (69.7%–75.3%)	28% (24.7%–30.3%)	100% (n = 1,008)

Note: Cells report row percentages with 95% CIs in parentheses; Pearson's χ^2 test $p = 0.2$.

tentative evidence that the very interested are the most likely to view several profiles, but they are not very different from users with the two medium levels of interest.

Finally, in terms of reading the parties' written explanations for their issue positions, we again find evidence that the very interested are no less likely to be explorers than less interested users (Table 7). In fact, the very interested users were again the most likely group to explore this option, but the differences to the less interested user groups are minor and not statistically significant. This corroborates our conclusion that politically interested users are not merely checkers and that they are certainly no less likely than less interested users to use VAAs to explore potentially relevant VAA information. While politically interested voters are generally not motivated to use VAAs to find out whom to vote for, some do seek out nuanced VAA information that they can use along with other nuanced information sources to decide whom to vote for.

Table 7. Political interest and opening the parties' written explanations.

Political interest	Did not open the written explanations	Opened written explanations	Total
Not at all interested	87% (81%–94%)	13% (6%–19%)	100% (n = 95)
Not very interested	90% (87%–94%)	10% (6%–13%)	100% (n = 336)
Somewhat interested	87% (83%–90%)	13% (10%–17%)	100% (n = 416)
Very interested	86% (81%–92%)	14% (8%–19%)	100% (n = 161)
All users	88% (86%–90%)	12% (10%–14%)	100% (n = 1,008)

Note: Cells report row percentages with 95% CIs in parentheses; Pearson's χ^2 test $p = 0.4$.

4.2.2. Does Vote Certainty Predict How Users Interact with VAAs?

van de Pol et al. (2014) suggest that politically undecided voters generally use VAAs to find out whom to vote for. Our results corroborate this expectation in the sense that undecided and uncertain users are much more likely to report this reason for using VAAs than users who have a clearer expectation about whom they will vote for (Table 8). In this sense, they can be characterized as explorers, but they are less likely to report using VAAs to learn about parties' opinions. This suggests an "exploring" type of behavior among the undecided, where they mainly rely on the advice to inform their vote choice. This is consistent with existing evidence that these users are influenced more by voting advice than users who have a clearer vote choice expectation (Germann et al., 2023; Tromborg & Albertsen, 2023).

Table 8. Vote choice certainty and self-reported motivations for using VAAs.

Pre-vote certainty	Learn about parties' opinions	Check if my party agrees with me	Find out who/what to vote for	Interesting test to reflect/discuss	Total
Undecided (No pre-vote party)	25% (18%–31%)	19% (13%–26%)	46% (38%–54%)	11% (6%–16%)	100% (n = 151)
Not at all certain (Has a pre-vote party)	20% (12%–28%)	39% (29%–49%)	36% (26%–46%)	5% (1%–9%)	100% (n = 100)
Quite certain (Has a pre-vote party)	27% (22%–31%)	39% (34%–44%)	22% (17%–26%)	13% (9%–16%)	100% (n = 366)
Very certain (Has a pre-vote party)	34% (29%–39%)	38% (32%–43%)	13% (10%–17%)	15% (11%–19%)	100% (n = 304)
All users*	28% (25%–31%)	36% (32%–39%)	24% (22%–27%)	12% (10%–14%)	100% (n = 921)

Notes: This table only uses data from those users who intended to vote for a party or did not know what party to vote for (N = 921). Pearson's χ^2 test $p < 0.001$.

However, when we zoom in on undecided voters' actual VAA behavior, the picture is more nuanced. Table 9 shows how vote choice certainty is related to the number of party profiles opened. The results show that undecided and uncertain users are the most likely to explore several party profiles as well. Especially those that have an intended party choice but are not at all certain are nearly twice as likely as those who are very certain about their vote choice (41% versus 23%). If uncertain and undecided voters also factor this information about non-recommended parties into their vote decision, then previous research has likely underestimated the effect of VAA advice on these voters—even if the estimated effects are already quite large for this group.

Table 9. Vote choice certainty and opening one or several party profiles.

Pre-vote certainty	Opened one party profile	Opened several party profiles	Total
Undecided (No pre-vote party)	70% (63%–78%)	30% (22%–37%)	100% (n = 151)
Not at all certain (Has a pre-vote party)	59% (49%–69%)	41% (31%–51%)	100% (n = 100)
Quite certain (Has a pre-vote party)	73% (68%–78%)	27% (23%–32%)	100% (n = 366)
Very certain (Has a pre-vote party)	77% (72%–81%)	23% (19%–28%)	100% (n = 304)
All users*	72% (69%–75%)	28% (25%–31%)	100% (n = 921)

Notes: This table only uses data from those users who intended to vote for a party or did not know what party to vote for (N = 921). Pearson's χ^2 test $p = 0.007$.

Finally, we find that undecided voters without an intended party choice are no more likely to read parties' written explanations than those who are quite certain or very certain about their vote choice (Table 10). However, those who have some idea whom they will vote for—but are not at all certain—are more likely to open written explanations than the remaining groups. Approximately one in five from this group engaged with the written explanations, while it was one in eight among other respondents. These numbers are likely conservative estimates of real-world behavior because they rely on paid survey respondents with no

incentive to engage more with the VAA than necessary. In sum, our findings for uncertain and undecided voters corroborate van de Pol et al.'s (2014) result that they tend to use VAAs to find out whom to vote for and suggest that they do not merely rely on VAA advice but also explore more nuanced VAA information about their party options.

Table 10. Vote choice certainty and opening parties' own written explanations.

Pre-vote certainty	Did not open the written explanations	Opened written explanations	Total
Undecided (No pre-vote party)	89% (84%–94%)	11% (6%–16%)	100% (n = 151)
Not at all certain (Has a pre-vote party)	82% (74%–90%)	18% (10%–26%)	100% (n = 100)
Quite certain (Has a pre-vote party)	90% (87%–94%)	10% (7%–13%)	100% (n = 366)
Very certain (Has a pre-vote party)	88% (85%–92%)	12% (8%–16%)	100% (n = 304)
All users*	88% (86%–91%)	12% (9%–14%)	100% (n = 921)

Notes: This table only uses data from those users who intended to vote for a party or did not know what party to vote for (N = 921); Pearson's χ^2 test $p = 0.14$.

5. Conclusion

In this article, we present novel behavioral data on how users interact with VAAs. We corroborate previous findings that many politically interested voters are motivated to use VAAs to check that they agree with their party or for entertainment and that undecided voters frequently use VAAs to find out whom to vote for. However, when we look at these users' actual behavior, we also find that a substantial share of interested, undecided, and uncertain voters also use VAAs in a more time-consuming way to explore the detailed answers of the parties they are able to vote for. This suggests that many VAA users in these groups are not merely checkers and doubters, but also explorers who obtain nuanced information from VAAs by using their additional features.

While this article comes a long way towards opening the black box of understanding users' VAA behavior, there are important ways in which future research can further strengthen our knowledge in this area. First, the sample is not restricted to voters who seek out VAAs in their own time. Therefore, our results can be used to infer how the average voting-age citizen interacts with VAAs, but we cannot know exactly how these results translate to the subset of voters who actually use VAAs on their own initiative. Second, and related, we ask participants to use a VAA and require them to spend some time on the results. It would be highly useful to track the exploring behavior of voters when they encounter VAAs in their own time, although this may be challenging in terms of obtaining participant consent and avoiding new sample selection issues. Third, we have only explored some VAA features, but there are many others that we have not assessed. For example, some VAAs use the VAA statement responses to graph the user's ideological placement in a two-dimensional space along with those of their candidates or parties. We do not know whether and how users engage with this type of information. Finally, future studies could fruitfully combine the kind of behavioral data collected

for this study with causal estimates of how VAAs affected users' vote choice. This would be a strong next step in linking how voters interact with VAAs with how they are affected by them.

In conclusion, the findings have important implications for both researchers and practitioners. By mapping out how users interact with VAAs, our findings support the idea that VAAs can serve multiple different informational roles for different types of voters. This has implications for theories on the role that information tools such as VAAs play in democracies. Moreover, our findings that voters with different levels of political interest and uncertainty are different in terms of how they approach and interact with VAAs have implications for theories on why VAA advice has heterogeneous effects. The findings are in line with prior research (Germann et al., 2023; Tromborg & Albertsen, 2023) and suggest that previous results may underestimate the true effects of VAA use on politically interested and undecided or uncertain voters. Finally, for designers and publishers of VAAs, our findings suggest that there is an audience for nuanced VAA features, such as the opportunity to explore all the parties' positions, because many voters actually interact with these options. Therefore, VAAs designers should take into consideration that there are different types of VAAs users, as there are voters with varying interests and motivations. Providing a streamlined user experience with limited user interactivity might not be the way to serve the growing and heterogeneous user base of these tools. Finally, low-interest users are more likely to be influenced by VAA advice, but they engage less with the more nuanced VAA functions (especially the least interested users). Consequently, it is critical that the initial advice is as methodologically sound, visually well-presented, and comprehensive as possible.

Acknowledgments

We thank our student assistants, Dagmar Vibe Huus Theilgaard and Marie Lindgaard Sundwall, for their excellent research assistance.

Funding

The research was funded by Independent Research Fund Denmark (Grant number DFF-2033-00003B) and by Koneen Säätiö (Kone Foundation).

Conflict of Interests

In this article, editorial decisions were undertaken by Diego Garzia (University of Lausanne/Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), and Andreas Albertsen (Aarhus University).

Data Availability

Replication materials will be made available on the corresponding author's website: www.mathiastromborg.com/articles

LLMs Disclosure

ChatGPT was used to check the grammar and spelling of some paragraphs.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Bundeszentrale für politische Bildung. (2021). 2002–2021: Die Geschichte des Wahl-O-Mat. <https://www.bpb.de/politik/wahlen/wahl-o-mat/326661/die-geschichte-des-wahl-o-mat>
- Dieing, T. I. (2025). Just can't get enough—profiling users of multiple voting advice applications. *Journal of Information Technology & Politics*. Advance online publication. <https://doi.org/10.1080/19331681.2025.2496259>
- Enyedi, Z. (2016). The influence of voting advice applications on preferences, loyalties and turnout: An experimental study. *Political Studies*, 64(4), 1000–1015.
- Garry, J., Tilley, J., Matthews, N., Mendez, F., & Wheatley, J. (2019). Does receiving advice from voter advice applications (VAAs) affect public opinion in deeply divided societies? Evidence from a field experiment in Northern Ireland. *Party Politics*, 25(6), 854–861.
- Garzia, D., & Ferreira da Silva, F. (2026). How many people use voting advice applications? Survey evidence from 15 democracies worldwide. *Politics and Governance*, 14, Article 11347.
- Germann, M., Fernando, M., & Gemenis, K. (2023). Do voting advice applications affect party preferences? Evidence from field experiments in five European countries. *Political Communication*, 40(5), 596–614.
- Hansen, K. M., & Stubager, R. (2019). *Den Danske Valgundersøgelse 2019*. <https://digidata.rigsarkivet.dk/aflevering/50084>
- Mahéo, V. A. (2016). The impact of voting advice applications on electoral preferences: A field experiment in the 2014 Quebec election. *Policy & Internet*, 8(4), 391–411.
- Munzert, S., Barberá, P., Guess, A., & Yang, J. (2020). Do online voter guides empower citizens? Evidence from a field experiment with digital trace data. *Public Opinion Quarterly*, 84(3), 675–698.
- Pianzola, J., Trechsel, A. H., Vassil, K., Schwerdt, G., & Alvarez, R. M. (2019). The impact of personalized information on vote intention: Evidence from a randomized field experiment. *The Journal of Politics*, 81(3), 833–847.
- Statistics Finland. (2023). Yli puolet suomalaisista käytti vaalikonetta ennen 2023 eduskuntavaaleja. <https://stat.fi/julkaisu/clpwcpcboal1760bvvs7ufutru>
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15(4), 582–599.
- 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.
- 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.

About the Authors



Andreas Videbæk Jensen is an assistant professor in political science at Aarhus University. His work explores how citizens participate in elections and which factors improve their conditions for doing so.



Veikko Isotalo is a PhD candidate in political science at the University of Helsinki, Finland. His research focuses on the effects of ideological positions on political candidates' electoral success in multi-party open-list electoral systems and on voting advice applications.



Mathias Wessel Tromborg is an associate professor in political science at Aarhus University. His research focuses on linkages between voters and their representatives in parliamentary democracies.

ARTICLE

Open Access Journal 

Enhancing Voting Advice Applications: Politicians' Perspectives on Additional Contextual Information and AI Integration

Elke van Veggel , Naomi Kamoen , and Christine Liebrecht 

Department of Communication and Cognition, Tilburg University, The Netherlands

Correspondence: Naomi Kamoen (n.kamoen@tilburguniversity.edu)

Submitted: 14 August 2025 **Accepted:** 27 October 2025 **Published:** 8 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Despite the popularity of voting advice applications (VAAs), users often struggle to fully understand the political statements. Since users rarely resolve these comprehension issues by searching for information outside the tool, one promising solution is to enrich VAAs with additional contextual information, either through clickable explanations in a so-called VAA+, or through a conversational agent VAA with an integrated chatbot. Responding to user-centric and normative calls for additional information in VAAs, the current study investigates how to add this information in a neutral and ethically sound way. In 20 semi-structured interviews with local politicians from two large Dutch municipalities, we explored (a) perceived acceptability and feasibility of four different types of contextual information—semantic clarifications, status quo descriptions, summaries of arguments in the political debate, and party positions—and (b) how AI could be used to disseminate this information. Discussions addressed appropriate resources, language use, and ethical concerns such as the risk of political bias. Findings show broad support for the addition of all four types of information, including summaries of pro and con arguments. Roughly half of the politicians emphasised that VAA developers should not evaluate the quality of arguments but could instead summarise the arguments as given by political parties to create concise and balanced overviews. Most politicians emphasised that implementation of AI is possible to some extent, but information must be accurate, politically neutral, and transparently sourced. This article reflects on the implications for theory and practice of future VAA development.

Keywords

artificial intelligence; ethical concerns; local politicians; voting advice applications

1. Introduction

Voting advice applications (VAAs) are online tools wherein users receive a voting advice based on their answers to a set of political attitude statements. These tools originated in the Netherlands as pencil-and-paper tools for high school students in 1989 and went digital in 1994 (De Graaf, 2010). Since then, they have become widely used across multi-party democracies (Garzia & Marschall, 2012). For example, the Dutch StemWijzer was used 9.1 million times during the 2023 parliamentary elections (“Recordaantal gebruikers voor StemWijzer,” 2023), and Germany’s Wahl-O-Mat saw 26 million uses ahead of the 2025 elections (Bundeszentrale für politische Bildung, 2025). While not all visits reflect unique users, these figures nevertheless show that VAAs reach a broad audience.

VAA brands work independently from political parties, and the tools are designed to be politically neutral (Garzia & Marschall, 2012). They aim to improve civic education, understood as the development of beliefs, skills, behaviours, and commitments for informed participation in society (Crittenden & Levine, 2018). Civic education involves more than developing policy preferences; it also includes fostering an understanding of why certain issues matter and how they relate to broader societal dynamics (Fossen & Anderson, 2014). To support this broader educational goal, some VAA brands include supplementary information alongside political attitude statements. This may involve definitions of key terms, contextual information about political issues, or explanations of party positions. Additionally, academics have started to explore information provision through the inclusion of chatbots using AI to deliver additional content (e.g., Kamoen & Liebrecht, 2022; Kamoen et al., 2024). The positive user evaluations of these enhanced tools indicate the need for additional contextual information in VAAs. Also, from a normative perspective, scholars explore the possibility of an effective VAA design with additional contextual information that encourages opinion reflection (Fossen & Anderson, 2014).

While supplementary information can foster civic education (Fossen & Anderson, 2014) and improve users’ understanding (Kamoen & Holleman, 2017), it also raises key questions: What types of information are ethically appropriate to include? And what role might AI play in generating and presenting this content? The current article explores these questions through interviews with local Dutch politicians. Although VAA providers operate independently, political parties are closely involved in the development of municipal VAAs. In StemWijzer, parties help select political statements and write their party positions; in KiesKompas, they direct developers to sources where official positions can be found (Gemenis & Van Ham, 2014). As such, politicians are key stakeholders in discussions about information provision and the potential role of AI in VAAs.

2. Theoretical Framework

2.1. Information Provision in VAAs

VAAs seek to enhance civic education in three ways: first, by procuring information about parties’ positions on policy issues; second, by analysing this information as to how well it fits with the users’ preferences; and third, by assisting voters in reviewing this information by providing matches with party stances at the end (Garzia, 2010). Empirical studies show that VAAs indeed positively affect outcome measures related to civic education, as they increase users’ perceived political knowledge (Van de Pol, 2016) and factual political

knowledge (Schultze, 2014), especially regarding parties' issue positions (Munzert et al., 2021). In addition, VAAs have been shown to boost voter turnout, although not consistently across studies and varying per research design and statistical method used for analysis (Munzert & Ramirez-Ruiz, 2021). Finally, VAAs affect vote choice, but the effect depends on the alignment with voters' existing party preferences (Enyedi, 2016), as well as on user characteristics such as degree of political interest (Tromborg & Albertsen, 2023).

Although VAAs contribute to civic education to some extent, there are also concerns about their effectiveness. One critique is that VAAs primarily respond to users applying an issue voting strategy when determining which political party to vote for (Fossen & Anderson, 2014; Fossen & Van den Brink, 2015). Issue voting (Downs, 1957) refers to voters making decisions by comparing their policy preferences to those of political parties, rather than basing their choice on party loyalty or other aspects unrelated to policy. VAAs are explicitly designed around this model of issue voting, as users are asked to express their opinions on a predefined set of issues and are then matched with the party that most closely reflects those views (for a more elaborate discussion see Fossen & Anderson, 2014).

Issue voting is closely related to a social choice view of democracy (Arrow, 1963). Social choice theory views democratic decision-making as a process of aggregating individual preferences into collective policies. Fossen and Anderson (2014) challenge this view, arguing that democracy is more complex than aggregating voters' preferences into collective policies. They advocate for a deliberative view of democracy, in which a "good" democracy is not defined by the translation of preferences into votes, but by the extent to which voters are encouraged to reflect on and possibly reconsider their preferences through new information. From this perspective, civic education entails more than the formation of policy preferences; it requires an understanding of why certain issues matter and how they relate to a broader societal context. As Fossen and Anderson (2014) argue, enriching VAAs with general background information and information about the arguments used in the debate would support such deliberation.

User-centric studies support the idea that it may be beneficial to include additional information in VAAs. In a think-aloud study among 60 voters, Kamoen and Holleman (2017) showed that, when no additional information is included, VAA users frequently struggle to interpret VAA statements due to a lack of semantic knowledge (word meaning), pragmatic knowledge (status quo), and argumentative knowledge (arguments pro and con to the statement). For instance, when evaluating a VAA statement such as "Income tax should be increased," some users may not know what income tax is (semantic knowledge), how high it currently is (pragmatic knowledge), or why such a policy might be proposed (argumentative knowledge). While users could, in principle, open a new tab to seek and find the required information, users rarely make this effort. Instead, they resort to satisficing strategies (Krosnick, 1991), disproportionately frequently picking a middle-response or no-opinion answer, which in turn undermines the validity of the voting advice.

Not only do user-centric studies indicate that comprehension problems occur when no additional information is included in a VAA, they also show that including information is beneficial, as this leads to higher political knowledge scores and better tool evaluations (Kamoen & Liebrecht, 2022; Kamoen et al., 2024; Van Zanten & Boumans, 2023). Taken together, both theoretical arguments (Fossen & Anderson, 2014) and empirical findings (Kamoen & Liebrecht, 2022; Kamoen et al., 2024; Van Zanten & Boumans, 2023) suggest that enriching VAAs with semantic, pragmatic, and argumentative information could improve their value as a tool for civic education.

2.2. The (Potential) Role of AI

In the Dutch context, some VAA brands incorporate additional content into their tool. In StemWijzer, for example, users can click on specific terms within VAA statements to access semantic explanations of difficult words. Additionally, users can click a button called “Know more” to access contextual information about the statement topic. We will refer to a tool like the Dutch Stemwijzer as a VAA+: a VAA with additional information in the web environment. In a VAA+, a limited amount of extra information can be included, with the advantage of the information being the same for every user and allowing control over the information presented.

An alternative approach is to connect a chatbot to a VAA. A so-called conversational agent voting advice application (CAVAA) allows users to ask numerous questions about the political issues presented in the tool, to which the chatbot provides a tailored response. CAVAs have been tested in a series of recent scientific studies (Dieing, 2025; Hankel et al., 2025; Kamoen & Liebrecht, 2022; Van Zanten & Boumans, 2023; Zhu et al., 2025), but to the best of our knowledge, have not been launched in real-life elections. Benefits of information provision via a conversational agent (CA) include the ability to provide an unlimited amount of information on-demand, that is, tailored to user preferences (Gemenis, 2024).

Kamoen and Liebrecht (2022) and Kamoen et al. (2024) demonstrate experimentally that CAVAs have advantages over VAAs without additional information, as they lead to higher political knowledge and tool evaluation scores. In their study, the chatbot had both predefined buttons where users could click on, as well as an open chat field. To be able to respond to questions typed via the open chat, a manually trained model based on open-source software categorised common user queries and triggered pre-formulated responses from a database; researchers then manually wrote the responses. While this setup allowed for controlling the information provided, it also has some limitations, also reflected by Dieing (2025): (a) creating human-written responses is time-consuming and resource-intensive; (b) the chatbot relies on fixed rules and intents, which means it can only handle the topics and phrasings it was explicitly programmed for; and (c) to make intent recognition work accurately, developers must manually gather a wide range of keywords and phrasings for each possible question.

To address these drawbacks, Dieing (2025) developed two CAVAs, both powered by a large language model (LLM). LLMs, such as ChatGPT, are AI systems trained on a large amount of data to understand and generate natural language responses to free-text user messages (Yigci et al., 2025). Dieing’s first CAVAA was an open-domain CAVAA, with unrestricted use of an LLM for text generation and intent recognition. Second, in the semi-open-domain CAVAA, Dieing (2025) applied three restrictions by adding filter layers that blocked user queries (prompts) that contained rude language, that requested the CAVAA’s political opinion, and that asked for a vote recommendation, to which the CAVAA then responded with a predefined answer. Next to technical testing of both chatbots, the semi-open CAVAA was also evaluated by an expert survey of German VAA creators. Both the open and semi-open CAVAs were able to answer questions about semantics, the status quo of statements, summaries of arguments, and party positions. While the expert survey indicated the potential for LLMs in CAVAs, implementing effective filter layers requires further refinement.

While Dieing (2025) assessed the effectiveness through technical testing and expert evaluations, Zhu et al. (2025) assessed an LLM-based CAVAA’s effectiveness among 331 users in the run-up to the 2024 German European Parliament elections. Zhu et al. (2025) used an LLM-based open-domain CAVAA, developed with

GPT-4o, meaning that an LLM was used without restrictions for both intent recognition and generative purposes. Results show that users valued the conciseness and personalisation of the chatbot's answers in explaining complex political issues. Moreover, participants felt motivated to actively reflect on their perspectives and understanding of the statements, as the chatbot invited participants to engage in an open-ended discussion, although some participants felt overwhelmed by this option. Furthermore, participants seemed to have an awareness of risks related to GPT-4o's black box nature and neutrality of its answers. The greater part of participants did not mind this, as long as the answers were written neatly, while some indicated a preference for seeing references in the chatbot's answers.

As partially discussed by Dieing (2025) and Zhu et al. (2025), implementing an LLM—albeit an open-domain system or one with restrictions—raises several ethical and technical concerns, such as concerns about the correctness of information, political bias, controllability of the system, user privacy, transparency, and the effort related to training a model. First, using an LLM raises the question as to whether the information that the system provides is factually correct, as LLMs are known to hallucinate (Jin et al., 2025). Second, LLMs have been shown to exhibit systematic political biases, often towards left-wing or pro-environmental ideologies (Dieing, 2025; Hartmann et al., 2023). Given that VAAs are designed to be politically neutral, such biases could pose serious issues. Third, LLMs such as OpenAI's GPT-4o are “black boxes” because their training data are kept private, making it difficult to understand how the models form their answers and to have control over the data (OpenAI, 2023). A fourth concern is the lack of verifiability of the information through sources, either because a non-existent source is given with the answer, or no source at all (Safdar et al., 2024; Zhu et al., 2025). Lastly, the issue of user privacy is a concern. Users may unknowingly disclose sensitive personal information while interacting with an LLM, which may be stored in the LLM's memory (Wang et al., 2025).

All in all, considering these earlier studies, we may conclude that current CAVAA research leans toward two opposing ends of the spectrum: either generative AI is entirely excluded from the development of chatbot content (Kamoen & Liebrecht, 2022), or it is utilised in real-time interactions with end users (Zhu et al., 2025). Yet, a compelling alternative remains underexplored: a human-in-the-loop approach wherein an LLM is used, but humans have control over the data that the system is trained on and give to the user (Mosqueira-Rey et al., 2023). In such a hybrid model, AI can support human contributors in multiple ways, for instance, by drafting initial versions of texts based on a select collection of sources or summarising arguments directly from political parties' party programmes—while preserving human judgment as a critical checkpoint in the process. In the present study, politicians will reflect on these different ways of implementing AI in a VAA context, as well as on the possible biases this may trigger.

3. Method

3.1. Design

To assess what politicians consider relevant additional information provision in VAAs and how AI can be used for this purpose, 20 semi-structured interviews were conducted from the 3rd of December 2024 to the 28th of January 2025. By choosing a semi-structured interview method, there is a clear framework around the main themes, while allowing flexibility to go more in-depth into specific topics as they emerge in the interviews. The Ethics Committee of Tilburg University approved the study (REDC2024.76).

3.2. Interviewees

We invited all political parties represented in two municipal councils in the Netherlands to participate, as well as both municipal registrars. Of all 25 political parties invited, 18 were both willing and able to participate in the interview. The municipal registrars of both municipalities agreed to participate in an interview as well. In total, 20 interviews were conducted in which 26 interviewees were interviewed, as some political parties sent two representatives.

Regarding the background and experience of the interview candidates, 18 (69%) had more than five years of experience in politics, five between three and five years (19%), and three between one and two years (12%). The most mentioned roles were municipal councillor ($N = 12$), faction leader ($N = 5$), or a combination of both ($N = 5$). Fourteen interviews included interviewees with experience in the development of VAAs, either in the form of crafting party stances or by participating in the statement selection process.

3.3. Material and Procedure

Interviews were conducted in person at the town halls of the municipalities. Our team contained four interviewers in total, of whom a minimum of two were present in each interview. The interviews were recorded with the permission of the party members. The shortest interview was 50 minutes, and the longest interview was 108 minutes, with a mean interview time of 68 minutes ($SD = 14.77$).

We started the interviews by reading an information letter. After the interviewees provided consent, the recording was started, and we asked demographic questions. The interviews were conducted following a predefined interview protocol. The interview protocol was designed based on two main topics: (a) information provision in VAAs and (b) the (potential) role of AI in VAAs.

3.3.1. Information Provision in VAAs

Within the topic of information provision, the four commonly used information types from previous (CA)VAA research were discussed: semantic information about word meaning; pragmatic information about the status quo; party stances for each party; and summaries of arguments pro and con without party reference. Visual tools were used to illustrate all four types of information. Visual tools facilitate the generation of ideas about topics for which there is not yet a common ground between two conversation partners (Glegg, 2019).

To explore how politicians feel about including semantic information, party members were presented with an example VAA statement: “All enforcers in the municipality should wear a bodycam.” Enforcer—in Dutch known as the abbreviation “BOA”—was specified by the interviewers as a potentially difficult word for users. Interviewers asked interviewees how they would explain this term to a VAA user who might not understand it, and what elements would make their explanation useful.

Pragmatic information was introduced using two additional VAA statements that are relevant in a municipal context—one about a referendum and another about a car-free city. Party members were asked to explain the status quo of the issue in their municipality, assuming that a user might request this background information. Again, they had to specify what made their explanations useful.

Next, politicians reflected on including party stances in the VAA. Party stances are already commonly included in Dutch VAAs; depending on the VAA brand, political parties either formulate their own stance texts (StemWijzer), or provide justifications derived from verifiable sources such as party manifestos (KiesKompas). Another key difference is the placement of party stances: KiesKompas displays them alongside the results page with the voting advice, while StemWijzer presents them alongside each statement. Interviewees reflected on these approaches.

To the best of our knowledge, summaries of pro and con arguments without party reference are not included in current Dutch VAAs. In a study by Kamoen et al. (2024), this information was created by selecting the two to three most common arguments from party stances, without mentioning party names. In the interviews, we discussed this type of information by showing interviewees a printed screenshot of Kamoen et al.'s (2024) operationalisation. Interviewees reflected on whether such information should be included and how the arguments should be selected and presented.

3.3.2. AI in VAAs

The topic of AI was structured around the various possible ways of using AI in VAAs. First, we asked questions related to the human-in-the-loop approach: using AI for searching within party programmes for party stances and arguments in the debate, instead of manually composing them. Second, questions around AI for text generation and intent recognition were asked, based around six general ethical considerations: (a) correctness of information (Jin et al., 2025); (b) political bias (Dieing, 2025; Hartmann et al., 2023); (c) verifiability of the system's training data (OpenAI, 2023); (d) verifiability of information through sources (Safdar et al., 2024; Zhu et al., 2025); (e) user privacy (Wang et al., 2025); and (f) the costs in terms of time and money when manually developing an AI model versus using a pre-existing LLM—and the options in between. Lastly, we examined the extent to which party members supported the implementation of five variations of CAVAA that differ in their use of AI: a CAVAA with button-based interactions only; a CAVAA with intent recognition and pre-formulated responses; an LLM-CAVAA that generates combinations of prewritten answers; an LLM-CAVAA that draws exclusively from a predefined collection of sources; and an LLM-CAVAA with full AI-based text generation. The full interview protocol can be accessed in Dataverse (<https://doi.org/10.34894/WDCCQG>).

3.4. Coding Scheme and Coding Procedure

The interviews were recorded with a recorder and through Microsoft Teams for backup. Afterwards, they were manually transcribed and then coded by two of this article's authors. Since some interviews included more than one interviewee, the unit of analysis used in the remainder of this article will be the interview, rather than the individual interviewee. Interviews were coded in an iterative manner. We constructed an initial coding scheme based on the interview protocol. We test-coded the first interview separately and discussed codes afterwards to refine the coding scheme. Subsequently, we separately coded a set of four interviews in the same manner. This process was repeated for the remaining interviews, meaning three more iterative rounds of (separate) coding of four interviews and discussions were conducted, until the final coding scheme was completed. Lastly, we once more checked all interviews for codes that did not match, discussed codes, and collaboratively developed final codes. This elaborate coding process allowed for continuous improvement of the codebook, which developed inductively throughout the iterative process.

Supplementary File 1 shows the full coding scheme, including a description and example per category. Via Dataverse, an example of a coded interview is provided (<https://doi.org/10.34894/WDCCQG>). For most interview topics, we coded whether interviewees were either positive, negative, or mixed about the inclusion of a certain type of information or some specific usage of AI. Moreover, a separate code was provided if the category was not discussed. Each code was accompanied by a written explanation provided by the coder, explicitly referencing the segment of the interview on which the code was based.

4. Results

4.1. Information Provision in VAAs

4.1.1. Semantic and Pragmatic Information

For both semantic and pragmatic information, interviewees frequently emphasised the importance of tailoring content. This emerged in two main ways: adapting descriptions to the local context and adjusting the language level to ensure accessibility for a broad and diverse audience.

In the interviews, interviewees reflected on general semantic and pragmatic explanations that were based on government websites. However, a preference was shown for explanations tailored to the local context of municipal elections or the municipality, rather than a general definition found online. Specifically, it was indicated in 15 interviews (75%) that a general definition for a word is often inadequate for political topics in the context of a municipal VAA. This is exemplified in the following quote: “[The concept] needs to be translated to how it works in a municipality and to a text that you think is understandable for the majority of the people you come across on the street” (I1).

Relatedly, in the majority of interviews, it was expressed that using a general definition might lead to verbose text, possibly including governmental jargon, which might cause users to drop out: “People will drop out [of the tool] if there is this much information in the text....The more information you provide, the more people will drop out” (I9). Hence, for both reasons of correctness and clarity, it was highlighted that local resources should be used as a baseline for both semantic and pragmatic information in the tool.

Second, it was stressed that—in general—clear language is essential. In the majority of the interviews, party members expressed that explanations should not introduce new concepts, real-world examples should be provided, and visual elements like a map or picture can help to clarify information about specific concepts or areas. The following quote illustrates this:

Adding a map could help. I also think this would be good for people who are less skilled at reading, that they see a map and think: ah, this is the ring of the city, and everything within this ring is the inner part of the city. (I19)

For checking whether explanations are written at a language level that is understandable for the general user group, it was recommended to check a Dutch website designed to examine whether a text complies with the basic B1 level of the Dutch language. If a text is labelled as too difficult, tips are provided for simplification.

4.1.2. Party Stances

A preference for implementing party stances alongside each statement was mentioned in 14 interviews (70%). Several party members note that presenting the stances alongside the statements is advantageous for delivering the information when it is needed, rather than presenting the stances at the end, when the opportunity has passed and users may no longer recall the original statement. Moreover, party members indicated that party stances help users to interpret the VAA statements “from a certain political party’s angle” (I3), to check “if a political party is trying to talk its way into a popular stance” (I14), and to compare different party stances “[if] you are doubting between some other parties” (I10). In some interviews, the advantages for politicians were also highlighted “because you can tell your story in a way, and it actually provides the opportunity to campaign a little” (I12).

In four interviews (20%), differing viewpoints were discussed. Regarding displaying party stances with the statements, interviewees stress that this makes the VAA more accessible for users: “This is far more accessible than an overview at the end” (I1). On the other hand, interviewees highlight that this can steer users in a certain political direction: “You will fill it in based on the party you already want to vote for beforehand” (I9).

4.1.3. Pro and Con Arguments

In 13 interviews (65%), party members felt positive about the implementation of argumentative information. The most frequently mentioned reasons were about the positive aspect of the absence of party names: “It gives users more interpretation per statement, without steering them based on what you think you should find important” (I2), and “It often is a matter of knowing enough about a certain political topic to really understand the pros and cons, and only then can you make a substantive decision” (I20).

One of the downsides mentioned was that it might be unsuitable to present subjective information such as arguments, because “we live in times of populism where scientific research is also questioned by some. This might cause harm to the [perceived] objectivity of a VAA” (I19).

4.1.3.1. The Source of Pro and Con Arguments

Besides reflecting on the inclusion of arguments, interviewees provided their opinion on how pro and con arguments should be selected and presented. In the majority of interviews (80%), interviewees expressed a preference for basing the arguments on party stances for various reasons, for example about users feeling heard through arguments coming from parties. In a different interview question, interviewees were asked about basing pro and con arguments on independent experts; in only five interviews (25%), interviewees expressed a preference for this, mainly because “I consider a scientist as capable of providing arguments pro and con in a neutral way” (I17).

Party members differed in their stances on how the arguments should be selected from party stances. In more than half of interviews in which it was discussed, interviewees expressed negative sentiments about compiling lists of arguments most frequently mentioned by the parties, because “the most frequently mentioned argument is not necessarily the decisive argument” (I14) and “if it is based on the most mentioned arguments, then you will always end up with the average argument, and an argument from a

more extreme party will never be mentioned" (I2). Moreover, making a selection was seen as presenting a biased overview of arguments to the users that might steer them in a certain political direction:

To what degree do the arguments you don't show also influence people's voting advice or choice? People are presented with a set of information and decide based on that. The person who determines which information will be shown in the VAA may also indirectly influence vote choices, if that person leaves arguments out that are mentioned by certain parties. (I14)

In the remainder of the interviews, party members were positive towards selecting the most frequently mentioned arguments, because "this allows you to select the main line of reasoning, let's say 80% of the arguments. You make a selection and clarify [to the user] that it is 80% of the mentioned arguments, hence not all the arguments" (I3).

4.1.3.2. Emotional and Factual Arguments

A division was apparent in the interviews over whether to include emotional arguments in the pro and con arguments or not, with interviewees being divided roughly in half around this issue. On the one hand, answers indicate that any argument should be included as an argument, for example based on experiences, because "there might not be a nuisance [in a certain neighbourhood] but people might experience it as such...those are two different things. If parties provide those arguments, you should include them all as arguments" (I18). Moreover, by not including every argument in the list, the situation as depicted in the VAA might not reflect the daily reality of the political landscape: "If you keep it purely factual, then yes, your VAA is factual, but at the moment users get home and turn on the TV, they see party leaders talking based on emotional arguments instead of factual ones" (I11).

In the other half of interviews, however, it is expressed that "including all arguments is not of much use, as then they might also quickly become party-specific again. If an argument for example states [statement], everyone will know who it's from" (I12), and emotional arguments "can be painful for certain groups in society. I would stay out of emotional language. This is not suitable for argumentative information and should remain in the party stances [i.e., with party names present]" (I6).

4.1.3.3. Presentation of Arguments

Once the arguments are selected, how should they be presented in a VAA? In the interviews where this was discussed, half of the participants were positive towards displaying all arguments, and the other half were not, or not without added considerations. The common thread, as shown in the following quote, is that there is a positive tendency towards giving users the opportunity to check all arguments, without presenting them all at once:

A solution is to, at first glance, show a selection of the most frequent arguments, with an option to click on a dropdown menu to show all arguments. As a user, you want the context, but you don't want to be flooded with information. (I1)

Arguments should not be presented exhaustively, but rather “short and clear. Do not add too much, because then people will not read it” (I16) and, in the case of the dropdown menu, it was recommended to “only add the option of the drop-down menu if the additional arguments really add something to the previous ones and are not just repetitions in slightly different words” (I8).

The above quotes suggest that participants prefer to visually display an equal number of the most frequently mentioned arguments at first glance, while making it optional for users to explore all arguments. In more than half of the interviews where it was discussed, party members indicated that the presentation of an equal number of arguments for and against would be useful to the user, because if users for example see more arguments pro than con for a statement, “they subconsciously have a bias towards ‘this is more beneficial to me than it is not’ [or vice versa]” (I20). Moreover, to not play into this bias, it was recommended that there should be equal amounts of lines of text as well:

I notice that in your example, there are three arguments for and against, but there are four versus three lines of text. This alone can visually influence users, and therefore, they might think that there are more arguments for this statement than against. (I4)

In some interviews, it was indicated that the quality of arguments is more important than the number of arguments displayed: “I find it more important that the arguments are based on fact than the number of arguments displayed” (I5), and “not every argument is equal to another [counter]argument” (I13).

4.2. AI in VAAs

4.2.1. Using AI to Search for Party Stances

AI can be used as a “search engine” for party stances in political parties’ programmes, instead of parties having to deliver stances to the VAA developer themselves. In 13 interviews (65%), however, party members stress the importance of writing and delivering the party stances themselves, as “not every party has written a stance in their party programme on each political statement [in the VAA], or they might have used different wording for it” (I10). In only four of the 20 interviews (20%), party members indicated that AI could be used for this purpose, as it can save them time in the VAA development process. In some interviews, it was, however, mentioned that “political parties should always have the ability to check whether the information is correct” (I8).

4.2.2. Using Generative AI

Instead of having humans write the additional information in a (CA)VAA, AI can be used for text generation. However, in 16 interviews (80%), party members indicated not being in favour of using an LLM for generative purposes in a VAA, mainly because of ethical considerations. Most prevalent in parties’ scepticism towards generative AI were concerns about the correctness of information. This was explicitly expressed in all interviews, for example, in Interview 5:

If there is only one wrong answer, you should already ask yourself if this is the right tool to implement, especially because this [the VAA] gives such heavy-handed advice to people...[and thus] people do not

fact-check information anymore. The [VAA] is an authority; people believe the information in there to be true. Hence you have to take that responsibility into account when developing this tool. (I5)

Another main ethical concern was the transparency regarding sources. In 13 interviews (65%), party members indicated the importance of presenting sources with the system's answers: "I find traceability of sources important, because then people can verify the information for themselves. Verifiability is good for the credibility of the VAA" (I9). Sometimes interviewees expressed more specific ideas about how sources can be integrated: "If an LLM searches the party programmes for the term 'safety,' it provides the passages to the user where safety is explained" (I2). Alternatively, another interviewee suggested providing hyperlinks: "Provide a hyperlink. It needs to be searchable, but not every user should be presented with all this information" (I17).

Moreover, related to the consideration of transparency about sources, is the consideration of the lack of controllability (i.e., black box nature) of LLMs. In eight interviews (40%), party members expressed concerns about a black box system such as GPT and using this for generative purposes:

By using GPT for text generation, you are dependent on ChatGPT and the sources they used for it. I don't know what sources they used. I do know they are a few years behind. I would find it very risky because you cannot control that. Also, because the servers are in America, a different territory. (I20)

4.2.3. Using AI for Intent Recognition

If (CA)VAA users are allowed to type in their information requests themselves, AI can be used for intent recognition. Interviewees mainly addressed ethical considerations regarding user privacy related to using AI for intent recognition. This was discussed in 16 interviews (80%). In the majority of those interviews, participants indicated not regarding user privacy as an issue: "I do not find user privacy more or less of an issue than in the case of existing VAAs. I assume that it's already being handled well" (I5) and "privacy is not a problem, but as a user, you have to be aware of the information you give to such a system. You should never talk about personal details" (I15). In some interviews, party members suggested providing the user with an opt-in option to decide whether the system trains itself further on their messages.

Moreover, a privacy issue that was explicitly mentioned, not only related to user messages but rather to all data the system contains, was the fear of hacking, data leaks, or misinformation distributed through unreliable sources: "I fear for data leaks. Answers from and about [municipality] can be used by secret services, should they have the intention to stir unrest in [municipality]. The data should remain ours" (I4) and "ChatGPT can be hacked. What if because of this a specific party is always presented as the best-matching party on the voting advice page, since it might be in favour of the hacker? Then you are in trouble" (I10).

4.2.4. Reflections on Specific Types of CAVAAs

The most restricted way in terms of AI implementation is to develop a CAVAA with only buttons, that is, one in which a chatbot is shown but no AI is used to recognise intent or generate texts. Such a CAVAA provides answers written by humans. We see that, in 18 interviews (90%), party members indicate being positive towards a CAVAA with buttons only. In two interviews (10%), party members explicitly indicated

not finding a button-based CAVAA to be sufficient. The following answer shows a positive response to a button-based CAVAA:

If you place a CA in a VAA, a lot of value is going to be placed on it [the CA]. Hence, you need to limit [the information] to information that is as factual as possible. This means that a CA[VAA] with buttons could work well. (I11)

In several interviews, the theme of correctness of information is again mentioned as an argument for a CAVAA with buttons only, as “with that, you decrease the complexity and limit the chance for mistakes” (I15) and “if you want the information to be trustworthy and controllable, a system with buttons works better. Maybe you should look for other ways to answer open questions” (I9). In some interviews, the use of a button-based chatbot was favoured, “as long as there will be enough additional information available through those buttons, such that it would be an actual improvement onto existing VAAs” (I8).

If the choice is made to let users type in their questions, as is currently the case in most CAVAA research, then the strictest model is to use AI for intent recognition and only provide answers written by humans. In most interviews where it was discussed (70%), party members answered generally positively to a CAVAA where AI is used only for intent recognition: “It can be a flatter system in which AI is used for intent recognition, because verifiability and trustworthiness are very important” (I10) and “The information is trustworthy, so that’s good, but it means you can ask fewer questions. Still, I think everything that helps to offer information in a low threshold manner is good” (I9).

In some interviews, however, a system with preformulated responses is criticised as being of little added value to current VAAs: “I can imagine it being useful, but then I wonder if it’s not already enough to just write better [static] explanations. ‘How big will the added value be?’ That’s what I’m wondering” (I19).

Instead of static answers prewritten by humans, a different option would be to implement an LLM for (a form of) text generation. This can be done in three ways: a complete implementation of an LLM, like for example ChatGPT; an implementation of an LLM that draws from a predefined collection of sources; or an LLM that is used for making combinations of prewritten texts. From all the interviews, we gather little support for a complete implementation of an LLM. Specifically, in 16 interviews (80%), party members are explicitly against giving full control to an LLM, expressing that “an AI system is not currently possible” (I19), “it is important to remain in control [over the system] ourselves” (I10), and “the correctness of information is too important” (I5).

The second means of implementing an LLM—letting it draw from a preselected collection of sources—received positive responses in most interviews in which it was discussed. According to the interviewees, “it is possible to develop a good chatbot if you train it on political parties’ opinions and council members, party programmes, motions, and policy reports” (I5), and “this ensures that the system does not just hallucinate something” (I6). In the remaining interviews, party members expressed that “when you develop a chatbot like this, it will probably go well in most cases. But I’m worried about the last 5% of the cases in which it does not...as that will gain attention” (I20).

Lastly, in roughly half of the interviews, party members are positive towards allowing an LLM to make combinations from predefined texts to better tailor the answers to the users’ questions: “I think this is a

good system. It seems more efficient, and answers do not contain that much information. The system should be able to answer as many questions as concisely as possible" (I7). In the other half of the interviews, party members expressed mostly negative views towards this implementation, such as: "It is hard to check if the combinations of the predefined texts are made correctly by the system. This way, you gather your information from a controlled environment, but is it then still accurate?" (I4)

5. Discussion and Conclusion

Adding supplementary information to a VAA can support civic education (Fossen & Anderson, 2014; Gemenis, 2024) and may increase users' understanding of the political issues behind the statements (Kamoen & Liebrecht, 2022). Our interview study offers insights into which types of information are acceptable to include according to politicians and explores the potential role of AI in generating and writing this additional information.

5.1. Semantic and Pragmatic Information

Interviewees made similar remarks on including semantic information about word meaning and pragmatic information describing the status quo: Writing such information requires customisation, in terms of content as well as language level. Interview candidates recommend focusing on local sources to write descriptions for municipal VAAs and to adjust the language level in descriptions to the B1 level of the Dutch language. The suggestion for accessible language use is also made in the expert evaluations of the LLM-based CAVAA in Dieing (2025). Interestingly, VAA brands already spend considerable effort in employing accessible language, for example with the help of linguists (ProDemos, 2025). However, a balance has to be established between using accessible language and ensuring the accuracy of information that conveys the complexity behind political decisions, which sometimes means that difficult words cannot be avoided. This nuance is also reflected in the answers of interviewees, who express a sense of disconnect between themselves and voters. Interviewees thus expressed a desire to create awareness among voters about what lies behind political decisions, while also wanting the VAA to be an accessible tool for all voters.

5.2. Party Stances

In real-life VAAs, the party stances on political statements are often either placed alongside each statement or alongside the results page with the voting advice. Most interviewees preferred displaying the party stances alongside the statements.

5.3. Pro and Con Arguments

Pro and con arguments are not included in current Dutch VAAs. Interviewees feel positively about the addition of this type of information. There is, however, no unanimous view on how to select and present the arguments. First, interviewees differed as to whether emotional arguments should be included in the arguments. On the one hand, it is seen as important to select only objective facts as arguments, as emotional arguments can hurt groups in society. On the other hand, emotional arguments are experienced by part of the electorate, and it would be unethical to keep this information from users.

Second, the interviews highlighted a discussion about the extent to which information that is (not) included in the arguments influences VAA users, as a certain selection of arguments may affect users' opinions on what arguments are important. Additionally, it was highlighted that arguments are subjective, and a selection made by researchers from a list of arguments is also subjective, even if this selection is based on a list of substantiated quality criteria for assessing arguments. Interviewees indicated that if arguments are selected based on criteria, these should be communicated to VAA users.

For presentation of the argumentative information, the results show a preference for a combination of two layouts. Specifically, the presentation of an equal number of arguments for and against at first glance, with the possibility for an exhaustive list of arguments to click on through a dropdown menu. This combination reduces the visual bias that might quickly steer users more towards one answer or the other, a phenomenon that has been extensively researched in dual-processing theories such as the Elaboration Likelihood Model (Petty & Cacioppo, 1984). Petty and Cacioppo (1984) showed that the number of arguments for a message can function as a simple cue to accept the message, especially when people are not very involved in the message content.

To the best of our knowledge, pro and con arguments are not yet used in practice but have been researched in previous CAVAA research (Kamoen et al., 2024; Zhu et al., 2025). Kamoen et al. (2024), presenting the pros and cons with a button in their CAVAA, show that this information was requested relatively often compared to the other three information types, especially by university-educated participants. On the other hand, Zhu et al. (2025) encouraged users to engage in discussion by inviting them to ask any questions in the open text field about possible uncertainties regarding the elections. While some participants felt encouraged to reflect on their perspectives and were introduced to unfamiliar topics, other participants felt overwhelmed and did not know what to ask the chatbot.

Hence, we can argue that showing VAA users relevant political arguments aligns VAAs more closely with a deliberative view of democracy, prompting reflection on political issues (Fossen & Anderson, 2014). Moreover, providing such arguments increases users' awareness of the complexity of political issues. However, this information should be presented in a structured manner, possibly with buttons, to guide users towards relevant contextual information and to reduce unawareness of their political knowledge gaps. It might encourage users to reflect on their perspectives while minimising cognitive effort, compared to having to think of relevant questions or arguments by themselves (Zhu et al., 2025).

5.4. AI in VAAs

5.4.1. AI as a Search Engine for Party Programs

Most politicians prefer not to automate the process of writing party stances. While understandable from the parties' perspectives, this may disadvantage VAA users. Research shows that party positions sometimes fail to reflect parties' actual voting behaviour during a previously served term (Škop, 2010). Moreover, discrepancies often occur between parties' answers and their corresponding positions, specifically when a neutral response masks a clear stance in the corresponding positions (Dieing & Siefken, 2025). Since politicians still typically write their own party stances, both VAA developers and users should be cautious of the accuracy of parties' positions.

5.4.2. Ethical Considerations Surrounding AI

Regarding the use of AI for information provision in VAAs, correctness of information was the most extensively discussed ethical consideration. Politicians underscored the responsibility of developing and distributing information in a VAA, as this tool is used to inform the majority of the electorate before voting. Results of the expert evaluations of the LLM-based CAVAA researched in Dieing's (2025) study support this finding, showing that most experts find it important that information in political advisory tools is correct. In the case of Dieing's (2025) study, this meant that the chatbot did not express its own opinion about statements or parties and stimulated the participants to think for themselves about voting choices. In light of the current study, the importance of correctness of information was expressed through suggestions that all information disseminated through a VAA should undergo a human review. Information should be factually correct, verifiable and politically unbiased; at the very least, there should be a human-in-the-loop that examines the information provided by the system before it is disseminated to the users.

To underscore the importance of this finding, the results of Zhu et al.'s (2025) study on an LLM-based CAVAA tested among users indicated that even though users are aware of the limitations LLMs such as GPT-4o might bring, the majority of the participants still trusted the chatbot's answers to be correct. The majority of the users did not question or check the CAVAA's answers, which, according to them, described various perspectives accurately, partly because the chatbot maintained a professional tone and communicated appropriately. It can be argued that these findings emphasise the importance of correctness of information, because it implies that users might take a chatbot's answers at face value if it merely seems appropriate and neutral, in the sense that attention is given to various perspectives. This might cause users to wrongly assess factually incorrect information as true.

Another important ethical consideration that was discussed in most interviews is transparency in the system and in terms of sources communicated to the user. This is also seen to be of importance among VAA users (Zhu et al., 2025) as well as in expert evaluations (Dieing, 2025). Implementing a transparent system also ensures user privacy. Politicians did not evaluate user privacy as a big concern, and there are two possible explanations for that. First, the focus of the question about user privacy was on whether the system would train itself further on user input and not for example on the possibilities of data leaks of information in the case of hacking. Had the focus been on the latter issue, it might have been perceived as more of a concern. Second, politicians emphasised the importance of the VAA having a database which the developers are in full control of, hence it should not be linked to an LLM such as ChatGPT, for example. Because politicians might have had this type of system in mind when reflecting on the question about user privacy, they might not have perceived user privacy as such a significant issue.

Taking into account all considerations about implementing AI in VAAs, it is better to be safe than sorry when developing a CAVAA. A great deal of importance lies in the correctness and controllability of the VAA. Although an LLM-based CAVAA is technically feasible (Dieing, 2025; Zhu et al., 2025) and evaluated positively by users (Zhu et al., 2025), a fully LLM-based CAVAA still displays errors, for example in terms of political bias in its answers (Dieing, 2025). In addition, it is important that information in a VAA is factually true, because users are inclined to assess information as true if it is communicated appropriately (Zhu et al., 2025). Politicians assess a more limited CAVAA system with preformulated answers as the most suitable option currently.

Therefore, the conclusion of this interview study is to ethically provide additional information in a VAA and develop a CAVAA with limited capabilities in terms of LLM usage. This means that a CAVAA with preformulated answers could work appropriately, similar to previously developed and researched CAVAAs (e.g., Hankel et al., 2025; Kamoen & Liebrecht, 2022; Kamoen et al., 2024; Van Zanten & Boumans, 2023). Argumentative information is seen as valuable to add but should be implemented in a structured manner to provide users with the knowledge needed to reflect on their political perspectives. This structured way of additional information provision in VAAs that encourages reflection more closely aligns a VAA with the notion of a deliberative democracy (Fossen & Anderson, 2014) and, consequently, better improves civic education among users.

5.5. Limitations and Suggestions for Future Research

Although the current study revealed valuable insights on politicians' views and considerations on information provision and the usage of AI in VAAs, it should be noted that only local politicians from two Dutch municipalities participated. The results should therefore be viewed within this context and might differ between municipalities or other levels of government in the Netherlands—and even beyond the Netherlands. Additional research is needed for national or international elections to assess if results are applicable in a broader context in which the role of politicians in the development of VAAs may differ, and, furthermore, could involve other stakeholders.

A next step after this interview study would be to apply the results from the current study in an experimental context, to assess how users would evaluate the suggestions highlighted by politicians. For example, the influence of visual cues by displaying different quantities of arguments, pro and con, can be examined, as well as the provision of sources alongside the additional information. Additionally, it is relevant for future research to study how information provision in VAAs should be realised: via a CAVAA or a VAA+.

Lastly, given the rapid developments of LLMs, it is important to realise that LLM-based VAAs are being used in practice (Otto, 2023) and that users often seem to trust LLM-based CAVAAs (Zhu et al., 2025), while there are many ethical considerations at play for implementations of such LLM-based tools in practice. It is imperative to remain aware of these developments and the ethical considerations associated with them; people should know not to blindly trust in LLMs, and VAAs should continue to be considered as reliable information sources by the average voter. Simultaneously, VAA brands and researchers should continue to collaborate in experimenting with LLM implementations in VAAs, to prevent being outpaced by opportunistic actors whose values may not align with those guiding the development of VAAs.

Acknowledgments

We would like to thank Rieke van Lieshout for her assistance in conducting the interviews.

Funding

This study was supported by the University Fund of Tilburg University.

Publication

Publication of this article in open access was made possible through the institutional membership agreement between Tilburg University and Cogitatio Press.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The interview guide, coding scheme, an example of a coded interview, and the examples used in the interviews are available via <https://doi.org/10.34894/WDCCQG>

LLMs Disclosure

The authors used GPT-4o solely to rephrase sentences for conciseness and grammatical correctness.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Arrow, K. J. (1963). *Social choice and individual values*. Yale University Press.
- Bundeszentrale für politische Bildung. (2025). *Bundestagswahl 2025*. <https://www.bpb.de/themen/wahl-omat/bundestagswahl-2025>
- Crittenden, J., & Levine, P. (2018). Civic education. In E. N. Zalta & U. Nodelman (Eds.), *Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/Entries/civic-education>
- De Graaf, J. (2010). The irresistible rise of StemWijzer. In L. Cedroni & D. Garzia (Eds.), *Voting advice applications in Europe. The state-of-the-art* (pp. 35–46). Scriptaweb.
- Dieing, T. I. (2025). AI-Driven Dialogue: Leveraging Generative AI in Conversational Agent Voting Advice Applications (CAVAAs). In A. Følstad, S. Papadopoulos, T. Araujo, E. L-C. Law, E. Luger, S. Hobert, & P. B. Brandtzaeg (Eds.), *Chatbots and Human-Centered AI* (pp. 161–180). Springer, Cham.
- Dieing, T. I., & Siefken, L. (2025). Read between the lines—How linguistics can unearth parties' true voting advice application positions. *German Politics*. Advance online publication. <https://doi.org/10.1080/09644008.2025.2548281>
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2), 135–150.
- Enyedi, Z. (2016). The influence of voting advice applications on preferences, loyalties and turnout: An experimental study. *Political Studies*, 64(4), 1000–1015.
- Fossen, T., & Anderson, J. (2014). What's the point of voting advice applications? Competing perspectives on democracy and citizenship. *Electoral Studies*, 36, 244–251.
- Fossen, T., & Van den Brink, B. (2015). Electoral dioramas: On the problem of representation in voting advice applications. *Representation*, 51(3), 341–358.
- Garzia, D. (2010). The effects of VAAs on users' voting behavior: An overview. In L. Cedroni & D. Garzia (Eds.), *Voting advice applications in Europe. The state-of-the-art* (pp. 13–47). Scriptaweb.
- Garzia, D., & Marschall, S. (2012). Voting advice applications under review: The state of research. *International Journal of Electronic Governance*, 5(3/4), 203–222.
- Gemenis, K. (2024). Artificial intelligence and voting advice applications. *Frontiers in Political Science*, 6, Article 1286893.
- Gemenis, K., & Van Ham, C. (2014). Comparing methods for estimating parties' positions in voting advice applications. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates: Voting advice applications in comparative perspective* (pp. 33–48). ECPR Press.

- Glegg, S. M. (2019). Facilitating interviews in qualitative research with visual tools: A typology. *Qualitative Health Research*, 29(2), 301–310.
- Hankel, S., Liebrecht, C., & Kamoen, N. (2025). 'Hi chatbot, let's talk about politics!' Examining the impact of verbal anthropomorphism in conversational agent voting advice applications (CAVAAS) on higher and lower politically sophisticated users. *Interacting with Computers*, 37(4), 221–233.
- Hartmann, J., Schwenzow, J., & Witte, M. (2023). *The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation*. arXiv. <https://doi.org/10.48550/arXiv.2301.01768>
- Jin, H., Zhang, P., Wang, P., Luo, M., & Wang, H. (2025). *From hallucinations to jailbreaks: Rethinking the vulnerability of large foundation models*. arXiv. <https://doi.org/10.48550/arXiv.2505.24232>
- Kamoen, N., & Holleman, B. (2017). I don't get it. Response difficulties in answering political attitude statements in voting advice applications. *Survey Research Methods*, 11(2), 125–140.
- Kamoen, N., & Liebrecht, C. (2022). I need a CAVAA: How conversational agent voting advice applications (CAVAAs) affect users' political knowledge and tool experience. *Frontiers in Artificial Intelligence*, 5, Article 835505.
- Kamoen, N., Liebrecht, C., Van Lieshout, R., & Van Veggel, E. (2024, August 12–15). "Hi voter, I can assist you in filling out this voting advice application!" How chatbot Lina assists voters in informing themselves about elections [Paper presentation]. ECPR General Conference, Dublin, Ireland.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236.
- Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2023). Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*, 56(4), 3005–3054.
- Munzert, S., Barberá, P., Guess, A., & Yang, J. (2021). Do online voter guides empower citizens? *Public Opinion Quarterly*, 84(3), 675–698.
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706.
- OpenAI. (2023). *GPT-4 Technical Report*. arXiv. <https://doi.org/10.48550/arXiv.2303.08774>
- Otto, R. (2023, November 19). Stemadvies van chatbot: 'Mooie toepassing, maar nog onbetrouwbaar.' *NU.nl*. <https://www.nu.nl/tech/6290114/stemadvies-van-chatbot-mooie-toepassing-maar-nog-onbetrouwbaar.html>
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46(1), 69–81.
- ProDemos. (2025). *Veelgestelde vragen*. StemWijzer. <https://stemwijzer.nl/faq>
- Recordaantal gebruikers voor StemWijzer: 9,1 miljoen. (2023, November 23). *ProDemos*. <https://prodemos.nl/nieuws/recordaantal-gebruikers-voor-stemwijzer-91-miljoen>
- Safdar, M., Siddique, N., Gulzar, A., Yasin, H., & Khan, M. A. (2024). Does ChatGPT generate fake results? Challenges in retrieving content through ChatGPT. *Digital Library Perspectives*, 40(4), 668–680.
- Schultze, M. (2014). Effects of voting advice applications (VAAs) on political knowledge about party positions. *Policy & Internet*, 6(1), 46–68. <https://doi.org/10.1002/1944-2866.POI352>
- Škop, M. (2010). Are the voting advice applications (VAAs) telling the truth? Measuring VAAs' quality. Case study from the Czech Republic. In L. Cedroni & D. Garzia (Eds.), *Voting advice applications in Europe. The state-of-the-art* (pp. 199–216). Scriptaweb.
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15(4), 582–599.

- Van de Pol, J. (2016). *Voting wiser: The effect of voting advice applications on political understanding* [Unpublished doctoral dissertation]. University of Amsterdam.
- Van Zanten, N., & Boumans, R. (2023). Voting assistant chatbot for ncreasing voter turnout at local elections: An exploratory study. In A. Følstad, S. Papadopoulos, T. Araujo, E. L-C. Law, E. Luger, S. Hobert, & P. B. Brandtzaeg (Eds.), *Chatbots and human-centered AI* (pp. 3-22). Springer.
- Wang, B., He, W., He, P., Zeng, S., Xiang, Z., Xing, Y., & Tang, J. (2025). *Unveiling privacy risks in LLM agent memory*. arXiv. <https://doi.org/10.48550/arXiv.2502.13172>
- Yigci, D., Eryilmaz, M., Yetisen, A. K., Tasoglu, S., & Ozcan, A. (2025). Large language model-based chatbots in higher education. *Advanced Intelligent Systems*, 7(3), Article 2400429.
- Zhu, J., Kempermann, M., Cannanure, V. K., Hartland, A., Navarrete, R. M., Carteny, G., Braun, D., & Weber, I. (2025). Learn, explore and reflect by chatting: Understanding the value of an LLM-based voting advice application chatbot. *CUI '25: Proceedings of the 7th ACM Conference on Conversational User Interfaces* (pp. 1–19). Association for Computing Machinery.

About the Authors



Elke van Veggel worked as a junior researcher in the Department of Communication and Cognition at Tilburg University until September 2025. Her research focused on assessing the effects of different versions of voting advice applications with the aim of better informing (young) voters during election times.



Naomi Kamoen is an assistant professor at Tilburg University. Her current research focuses on the uses and effects of voting advice applications. She collaborates with societal partners in real-life election settings, using a wide range of research methods such as experiments, content analyses, and big data analyses.



Christine Liebrecht is an assistant professor at Tilburg University. Her research examines the transformation of human communication through technology across different user types and contexts. Her approach incorporates both quantitative and qualitative research methodologies, frequently linking insight to practical implications.

Estimating the Recommendation Certainty in Candidate-Based Voting Advice Applications

Fynn Bachmann , Daan van der Weijden , Cristina Sarasua , and Abraham Bernstein 

Department of Informatics, University of Zurich, Switzerland

Correspondence: Fynn Bachmann (fynn.bachmann@uzh.ch)

Submitted: 31 August 2025 **Accepted:** 12 November 2025 **Published:** 21 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Voting advice applications typically require users to answer questionnaires before receiving party or candidate recommendations. As users answer more questions, the recommendations naturally become more accurate. However, when users do not complete the questionnaire, the certainty of these recommendations is unknown. In this work, we develop and present a measure to quantify this certainty by introducing an algorithm that estimates the candidate recommendation accuracy—the overlap between early and final recommendations—after each question. Through simulations based on existing voter data, we find that our algorithm is more accurate than heuristic estimates. Additionally, it can identify *stable recommendations*—candidates who are likely to be among the final recommendations—with fewer false positives. Furthermore, we conduct a user experiment investigating different ways of communicating recommendation certainty to users. Our results show that users answer more questions when they see a preview of stable recommendations, but quit the questionnaire earlier when we display an artificially high candidate recommendation accuracy estimate. Moreover, we find that users appreciate the interface's simplicity over its accuracy. We conclude that displaying personalized stable recommendations can spark curiosity towards voting advice applications while providing a robust estimate of recommendation certainty for users who submit incomplete questionnaires.

Keywords

human-computer interaction; personalized interfaces; recommendation quality; recommender systems; statistical modelling; voting advice applications

1. Introduction

Voting advice applications (VAAs), lately described as “democratic recommender systems” (Berdoz et al., 2025), contribute to political education in multi-party democracies by recommending candidates or parties to voters (Garzia & Marschall, 2019; Tromborg & Albertsen, 2023). During elections, voters complete specifically designed questionnaires to identify their closest matches among a set of political actors (Louwerse & Rosema, 2014). In this way, VAAs inform citizens about pre-election party positions (Schwarz et al., 2010), which has been shown to positively affect turnout, vote choice, and issue knowledge (Garzia et al., 2017; Ladner et al., 2012; Munzert & Ramirez-Ruiz, 2021).

While the design of VAAs has been extensively studied (Bruinsma, 2020; Buryakov et al., 2024), there has been limited work published on users’ primary concerns: the *quality* of their recommendations. Existing work on recommendation quality mainly focuses on the algorithmic perspective, i.e., on the robustness of the matching function (Berdoz et al., 2025) or the impact of scales (Rosema & Louwerse, 2016). However, it remains unclear how accurate the recommendations are if users submit incomplete questionnaires. This is the case for approximately 76% of Smartvote users, the most popular VAA in Switzerland, where 34% of users answer less than 30 out of 75 questions. Bachmann et al. (2024) approached this topic by introducing an adaptive question selection mechanism to increase the accuracy for early dropouts. Still, they did not provide a method to estimate the certainty of early recommendations. The question of whether and how such an estimate influences users and what measure of certainty they would appreciate most seems to be a gap in the literature, especially in candidate-based VAAs, where the choice is significantly more complex.

The contributions of this article are two-fold: First, we define the quality of early recommendations as the overlap with the final recommendations and develop an algorithm that estimates this metric before all answers are known by predicting users’ remaining responses. Through simulations based on existing voter data obtained from Smartvote (Politools, 2023), we evaluate the precision of the resulting *recommendation certainty*. Second, we design and experimentally test three interfaces that communicate the certainty forecast to users. Inspired by Law et al. (2016), some of the interfaces are designed to trigger curiosity by displaying a preview of the most likely candidate recommendations. To evaluate these interfaces, we conduct a controlled experiment where users can finish the questionnaire when they feel they have answered enough questions. Together, these two contributions address the following research questions:

RQ1: How precisely can our algorithm estimate the recommendation certainty?

RQ2: How does the display of the recommendation certainty affect user behavior in candidate-based VAAs?

RQ3: Which interface works best to communicate the recommendation certainty comprehensively to users?

The results of the experiments demonstrate that our method can predict individual progressions of the recommendation certainty, capturing sudden changes in the response pattern of each user. Moreover, the preview of stable recommendations increased the number of questions that users answered before exiting the questionnaire, therefore enhancing engagement. Conversely, users dropped out earlier when the

displayed certainty was artificially high, indicating that users consider recommendation certainty when deciding whether to answer another question. These findings motivated us to develop *semantic progress indicators* in sequential questionnaires that estimate the level of completion for each user individually, rather than the typical static progress bar. User feedback collected in the post-survey revealed that users understand simpler interfaces better than more accurate ones—creating a trade-off for VAA developers when designing semantic progress indicators for VAAs in practice.

2. Related Work

VAAs are widely used in multi-party democracies (Garzia & Marschall, 2019; Germann & Mendez, 2016). In more than 10 European countries, at least 10% of the population regularly engage with such online tools before elections. The German Wahl-O-Mat has by far the largest audience, with almost 26 million uses in 2025. Controlling for the population size, Norway has the highest share of VAA users, with over 60% of voters consulting a VAA before the election in 2021. Underlining their importance, it has been shown across many studies that VAAs impact reported election turnout (Germann & Gemenis, 2019; Munzert & Ramirez-Ruiz, 2021).

2.1. User Studies in VAA Research

Several randomized field experiments were conducted to measure the effects of VAAs on affective polarization, political knowledge, or vote intention (Garzia et al., 2017; Pianzola et al., 2019). However, this recruiting approach often suffers from self-selection bias (Ladner et al., 2012). To the best of our knowledge, few controlled user experiments were implemented. Bruinsma (2020) investigated the effect of visualizations in VAAs and found that many users had difficulties understanding them, regardless of their political interest or graphical knowledge. Kamoen and Liebrecht (2022) ran controlled user experiments to evaluate their large language model-enhanced conversational agent VAA. Users who interacted with the conversational agent VAA reported higher political knowledge and answered more factual knowledge questions correctly. We contribute to the body of work by conducting a controlled user experiment that investigates the effect of visualizing recommendation certainty to VAA users (Kay et al., 2016).

2.2. VAAs as Democratic Recommender Systems

The effect of the matching algorithm in VAAs has been intensively explored. It was often found that the distance function to compute matches drastically affects recommendations (Berdoz et al., 2025; Louwerse & Rosema, 2014). While such distance functions are most frequently used, a different approach is to learn recommendations based on response patterns of other voters (Romero Moreno et al., 2022). Social VAAs integrate a network approach into the matching algorithm by learning preferences based on other users' recommendations (Agathokleous & Tsapatsoulis, 2016; Katakis et al., 2014). This connects VAA research to collaborative filtering (Elahi et al., 2016; Rubens et al., 2015) and neighborhood-based methods of recommender systems (Lü et al., 2012). Such systems are compelling in e-commerce, where they predict final preference profiles as responses trickle in. We use a statistical method from collaborative filtering to estimate the recommendation certainty in VAAs.

2.3. Statistical Models for Political Data

Political scientists have developed several statistical models that visualize candidates or voters in a low-dimensional space (Clinton et al., 2004; Poole & Rosenthal, 1985). These models take features (such as roll-call data, survey responses, or social media interactions) and identify the ideal points of both the political actors and the features. If an actor is embedded close to a feature, they are more likely to agree with it. However, missing values are a considerable problem for most of these models. Often, mean imputation is used for simplicity, while item-response theory frameworks yield more nuanced results, especially for political data (Bachmann et al., 2024). Iterative algorithms, such as iterative principal component analysis (Grung & Manne, 1998), multiple imputation by chained equations (Buuren, 2012), or variational autoencoders (McCoy et al., 2018), do not yield a single maximum-likelihood imputation. Instead, they treat missing entries as latent variables and iteratively update or sample them using approximate posterior distributions. In our case, we adopt an iterative procedure that samples personalized completions of partially filled questionnaires based on users' inferred positions in the political space.

3. Methods

We propose a novel algorithm to estimate the recommendation certainty throughout a candidate-based VAA questionnaire. The core of this method is a statistical model that can predict users' missing answers based on their responses to the initial questions. This is necessary to gauge users' final recommendations. In this section, we explain the details of the statistical model and describe the estimation algorithm in more detail.

3.1. Statistical Model

In the trade-off between expressiveness and interpretability, we chose a two-dimensional latent-space model where the dimensions are learned from the distribution of the training data, rather than being predefined by human experts. Specifically, we follow the approach of Potthoff (2018) and Bachmann et al. (2025), where the statistical model is a combination of principal component analysis (PCA) and logistic regression (LR). Based on PCA projection of the candidates' answers into the latent space, an LR is learned for each question. This training procedure results in two ideal-point parameters per candidate (x_1, x_2) and three LR parameters per question (β_1, β_2, α). The likelihood of answering "yes" to question q (given a position x in the latent space) is defined as:

$$\hat{y}_q(x) := P(y_q = 1 | x) = \sigma(x^T \beta_q + \alpha_q)$$

Where σ is the sigmoid function:

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

And $x^T \beta$ is the scalar product of the candidates' position and question parameters. To embed new users into the latent space, we can then use their given answers to find the optimal position $\tilde{x}$. This optimal position (maximum likelihood estimate) minimizes the distances of the real given answers y_Q and the models' predictions $\hat{y}_{q \in Q}(x)$ such that:

$$f(x; y_Q) := \prod_{q \in Q} (1 - |\hat{y}_q(x) - y_q|)$$

Is maximal at $f(\tilde{x})$. Normalizing this function $f(x; y_Q)$ with respect to x then leads to the posterior distribution:

$$g(x; y_Q) := P(x | y_Q) = \frac{f(x; y_Q) \mathcal{N}(x; \mu, \Sigma)}{\int f(x; y_Q) \mathcal{N}(x; \mu, \Sigma) dx}$$

Where we include a Gaussian prior $\mathcal{N}(x; \mu, \Sigma)$ with mean μ and covariance matrix Σ to avoid the user positions moving to infinity. Note that if the set of answers is incomplete, the missing values are ignored for the embedding. This ensures more nuanced embedding for sparse data.

3.1.1. Discretization

By embedding users into the latent space, we can predict all remaining questions for a user by combining their position and the questions' likelihoods as previously described. However, minimizing the posterior distribution to obtain the maximum likelihood estimate is not analytically solvable. We, therefore, discretize the space into a grid of 200×200 points. We choose this grid size because it is sufficiently expressive and reasonably small, allowing for efficient computation. Figure 9A in the Supplementary File shows, as an example, the distribution of candidates' positions in the latent space and visualizes the likelihood function.

3.1.2. Sampling

The posterior distribution $g(x; y_Q)$ indicates the latent ideology of the response pattern y_Q (see Figure 9 in the Supplementary File). However, given the Bayesian approach, each position has uncertainty. To predict missing values with variations (Buuren, 2012), we can draw samples of the posterior distribution instead of taking the maximum likelihood predictions defined by $\hat{y}(\tilde{x})$. Then, we predict the remaining answers for each of the sampled positions to obtain many possible completions.

3.2. Recommendation Certainty

In candidate-based VAAs, users receive as recommendations the set of k candidates with the smallest distance to their response profile y_q (k -nearest neighbors). This distance is computed using the L_1 norm (or "Manhattan" distance) of the user's and candidates' answers:

$$D_M(c, y_Q) = \sum_{q \in Q} |y_q - c_q|$$

If the number of answered questions is less than the total number of questions ($Q < 75$), the remaining candidates' answers are ignored. We refer to the k candidates that minimize this distance to the response pattern y_Q of the user as *early recommendations* R_Q . However, through the predictions of the statistical model, it is possible to impute the remaining answers of the user. If these predictions are included in the distance computation, we refer to them as *predicted recommendations* R_P . After completing all questions, users receive their *final recommendations* R_F . Note that $R_Q = R_P = R_F$ if $Q = 75$. Figure 1 provides an overview of these three recommendation types.

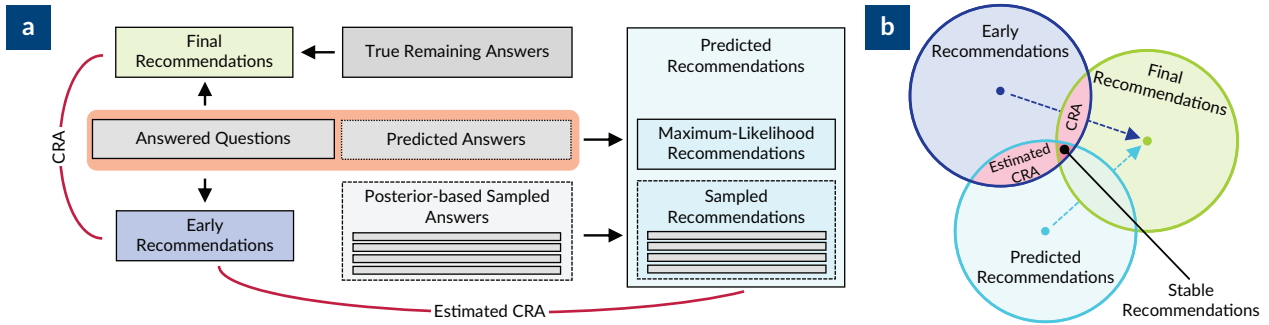


Figure 1. Schematic sketch of the recommendation types and the candidate recommendation accuracy (CRA) estimation algorithm. Notes: (a) The CRA is computed as the overlap of the *final* and *early recommendations*; it can be estimated using the *predicted recommendations*, which are informed by sampling the remaining answers; (b) At the beginning of the questionnaire, all types of recommendations are distinct sets without overlap; throughout the questionnaire, the early and predicted recommendations move towards the final recommendations; if the overlap of early and final recommendations increases (CRA), the overlap of early and predicted recommendations increases as well (estimated CRA); the overlap of all three types is called *stable recommendations*.

3.2.1. Candidate Recommendation Accuracy

The candidate recommendation accuracy (CRA) quantifies the quality of early recommendations R_Q as their overlap with the final recommendations R_F . More generally, the CRA of any set of candidates R is given by its overlap with R_F :

$$\omega(R; R_F) = \frac{|R \cap R_F|}{k}$$

where $k = |R_F|$ is the number of candidates per recommendation. Here, the final recommendations serve as “ground truth,” i.e., the optimal set of recommendations possible within the framework of a questionnaire-based VAA—even though they might not globally be the best set of candidates for a user.

3.2.2. Estimation Algorithms

Calculating the CRA is only possible *a posteriori*, i.e., when all questions are answered, and the final recommendations are known. To estimate the certainty of early recommendations *during* the questionnaire, we use the predicted recommendations as a proxy for the final recommendations. As shown in Figure 1a, we calculate the overlap between the early and predicted recommendations to estimate the CRA. This does not imply that the predicted recommendations are necessarily correct. As Figure 1b indicates, the early and predicted recommendations approach the final recommendations from different directions. Even if both are equally false, their overlap estimates *how false* both are.

We evaluate two approaches to predict the remaining answers for CRA estimation: In the first approach, we use the maximum likelihood predictions. We call this approach One-Shot. The estimated CRA is then given by $\omega(R_Q; R_P)$. In the second approach, we use sampling to generate variations in the predictions (Posterior). Here, the CRA estimate is given by the average overlap:

$$\frac{1}{M} \sum_m^M \omega(R_m; R_Q)$$

Where R_m is the set of recommended candidates obtained from the sample m , and ω is the CRA defined above. This estimate thus captures how similar the candidate recommendations are across different plausible completions of the user's answers. We compare both estimation algorithms to the baseline of user-agnostic estimations, such as a linearly increasing estimate (Linear) and aggregated CRA values from the Smartvote data (Historic). The Linear estimate corresponds to a static progress bar increasing from 0% to 100% in even steps, while the Historic estimate adjusts for the average CRA progression of existing user profiles (see Figure 2b). We chose these baselines as they are straightforward alternatives to displaying the recommendation certainty to users.

3.2.3. Stable Recommendations

Lastly, we define *stable recommendations* as those candidates in the early recommendations that will also appear in the final recommendations. They can be used in the preview of recommended candidates, where it is essential to display only a few false positives. We compare two approaches to select stable recommendations and evaluate them by the number of true and false positives: In the first approach, we select the overlap of the early and predicted recommendations. This selection of size $|R_p \cap R_Q|$ is initially empty and then grows as the early and predicted recommendations converge to the final ones. The advantage of these *estimated recommendations* is their fast computation with just one maximum-likelihood prediction. In the second approach, we use sampling to add variance to the predicted recommendations. Then, we compute the frequency with which each candidate appears in the *sampled recommendations* R_M . If this frequency exceeds a certain threshold, we select the candidate. A shortcoming of this approach is that the threshold must be learned from previous user interactions. However, it is convenient that the candidates' frequencies converge to the sampled CRA estimate when sufficiently increasing the sample size.

4. Evaluating the CRA Estimation Algorithms

In the first study, we simulate the progression of recommendation certainty for existing voters in the Smartvote data and evaluate the estimation algorithms based on their accuracy. In this section, we describe the data used, the framework, and the simulation results.

4.1. Data

The Smartvote data were obtained during the Swiss National Elections in 2023 (Politools, 2023). They include the answers of candidates and voters from Zurich—the largest canton in Switzerland in terms of population—to 75 political questions, recorded on a Likert scale with options ranging from 4 to 7. We convert these ordinal responses into a continuous scale from 0 (*complete disagreement*) to 1 (*complete agreement*), evenly spacing the remaining options across this range. In addition, the data include personal information about the candidates (e.g., party affiliation, mottos, budget, age). In total, we use the data of 1,029 Zurich-based candidates, where we focus on candidates from the eight major parties (represented in the Federal Assembly) and categorize candidates from smaller parties under “Others.” Detailed information on these parties can be found in Table 3 in the Supplementary File.

Additionally, to accurately evaluate the precision of the recommendation certainty, we generate a representative sample from the voters in the Smartvote data. Our sampling procedure involves iterating through each combination of gender, age group, and political position, selecting the appropriate number of voters to obtain 1,122 representative individuals. This approach ensures that our voter data accurately reflect the demographic and political landscape of the canton of Zurich and are comparable to the sample in our user experiments. We consulted demographic data from Urbistat (2025) and election results from the Swiss Federal Statistical Office (2023). Figure 16 in the Supplementary File shows the distribution of voters in the representative sample.

4.2. Simulation Framework

We use the representative voter sample from the Smartvote data to identify the best-performing estimation algorithm through a simulation. As described in Section 3.2.1, the final recommendations are based on all voters' answers as their k -nearest neighbors in the candidate's data. We consider that $k = 36$ as this is the number of candidates per recommendation in the canton of Zurich (for comparison, the results for $k = 10$ are provided in the Supplementary File in Table 4 and Figures 10 to 14). For each user, we iterate through the questions and compute the early and predicted recommendations after each answer. To predict the recommendations, we compare the different approaches explained in Section 3.2.2. We then compare the true CRA, the estimated CRA, and the set of stable recommendations (see Section 3.2.3). For the sampled recommendations, we consider thresholds $t \in \{0.1, 0.2, \dots, 0.9\}$. Together, this simulation addresses RQ1, exploring whether our algorithm can accurately forecast the recommendation certainty.

4.3. Simulation Results

The simulation was implemented in Python and run on a Mac with an M1 processor. Using the *sklearn* k -nearest-neighbor tree, all algorithms computed the recommendation certainty in less than 90 ms per estimate. The number of samples did not drastically impact this duration. For example, it took 18.1 ms to calculate the recommendation stability based on the maximum-likelihood estimate, while it only took 82.9 ms to do the same with 1,000 Posterior samples. The fastest methods were the user-agnostic estimates with less than 1 ms look-up time. This puts all times well below the threshold of 0.1 s for seeming instantaneous reactions (Card et al., 1991).

4.3.1. Error of the Estimated CRA

We tested four different approaches to estimate the CRA throughout the questionnaire. As seen in Figure 2a, the average error depends on the number of questions answered. All estimation algorithms start with an error of 4% which increases as users answer more questions. The best-performing approach, 1,000 Posterior samples, then reaches a plateau after seven questions with an error of 8.7%. For fewer samples, the error is slightly higher. In contrast, the maximum-likelihood approach, One-Shot, has a maximum error of 10.7% after seven questions. Still, this result is better than the baseline approaches, Linear and Historic, which have an average error of 13.8% and 11.5% respectively, at this stage.

Table 1 quantifies how accurately the algorithms estimate the true CRA averaged across all users and questions. The worst-performing approach, Linear, has an average error of 10.22%, while the

best-performing algorithm, 1,000 Posterior samples, is on average 6.28% off. Here, the number of samples significantly affects the performance. We find that more samples have smaller errors, fluctuate less, and require longer computation times. While the decrease of the error becomes very small for sample sizes larger than 100, the paired *t*-test still shows that they are significant ($p < 0.001$). Of all deterministic algorithms, the maximum likelihood estimate reaches the best results with an average error of 7.27%, which is significantly better than the baselines of Linear and Historic ($p < 0.001$).

Table 1. Fluctuation, error, and duration for different estimation algorithms.

Algorithm	Method	Fluctuation (%)	Error (%)	Duration (ms)	Recommendation
CRA	Ground truth	3.35 ± 0.02			Only a-posteriori
Historic	User-agnostic	1.25%	9.33 ± 0.12	< 1 ms	Requires user data
Linear	User-agnostic	1.33%	10.22 ± 0.14	< 1 ms	Linear progress bar
One-Shot	Maximum-likelihood	3.90 ± 0.02	7.27 ± 0.08	25.1 ± 0.3	Includes estimated recs.
Posterior	1 sample	7.44 ± 0.04	7.95 ± 0.07	25.9 ± 1.1	Includes sampled recs.
Posterior	10 samples	3.63 ± 0.02	6.47 ± 0.07	26.6 ± 0.7	Includes sampled recs.
Posterior	100 samples	2.88 ± 0.01	6.30 ± 0.07	31.6 ± 1.0	Includes sampled recs.
Posterior	1,000 samples	2.78 ± 0.01	6.28 ± 0.07	82.9 ± 4.1	Includes sampled recs.

Notes: The fluctuation is computed as the sum of absolute successive differences; a higher value means the estimations are less smooth; the error gives the mean absolute difference between the estimated CRA and the true CRA; the duration includes computing the estimated CRA and stable recommendations; all values are means and their standard errors.

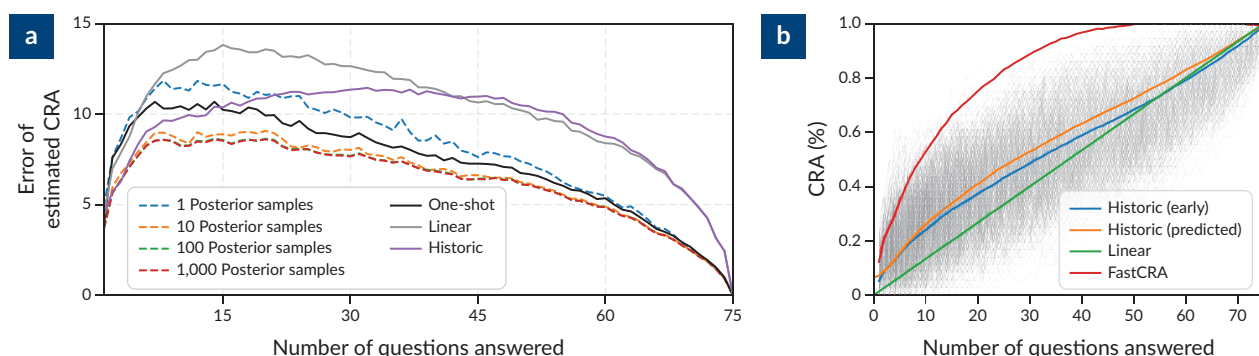


Figure 2. Mean absolute error of the estimation algorithms compared to the true CRA. Notes: (a) For all estimation algorithms, the error increases from 4% to around 7–12% after seven questions; the user-agnostic predictions, Linear and Historic, perform worse than Posterior and One-Shot; (b) Each voter has an individual progression of the true CRA (gray lines); the Historic estimates correspond to the average CRA across all voters.

4.3.2. Fluctuation of the Estimated CRA

Defined as the sum of absolute differences between the current and the previous estimated CRA, a high fluctuation indicates that the estimates strongly vary from question to question. However, as shown in Table 1, the true CRA also fluctuates with a mean value of 3.35%. A similar fluctuation is achieved by One-Shot with a value of 3.90%. Sampling reduces the fluctuation to 2.88% and 2.78% (for 100 and 1,000 samples, respectively). This difference is statistically significant with a *p*-value of $p < 0.001$ in a paired *t*-test.

The lowest fluctuation is achieved by user-agnostic estimates. Linear estimates achieve 1.33%, while the Historic achieves 1.25% (note that the initial estimate here is larger than 0%).

4.3.3. Precision of the Stable Recommendations

In Figure 3, we inspect the empirical relation between the Early and Predicted recommendations. Initially, these two sets are distinct from each other (and from the final recommendations). As users answer more questions, all sets increasingly overlap until they are finally identical. Indicated by the blue and orange lines, both Early and Predicted converge to the final recommendations at a similar pace. The number of stable recommendations (those candidates who are part of all three sets) monotonically increases from 0 to 36, as evidenced by the pink line. This indicates that even after 15 questions, 50% of the Estimated recommendations are stable. Already after 35 questions, 75% of them are true positives. We therefore find that the predictions uncover different but similarly accurate recommendations compared to the early ones.

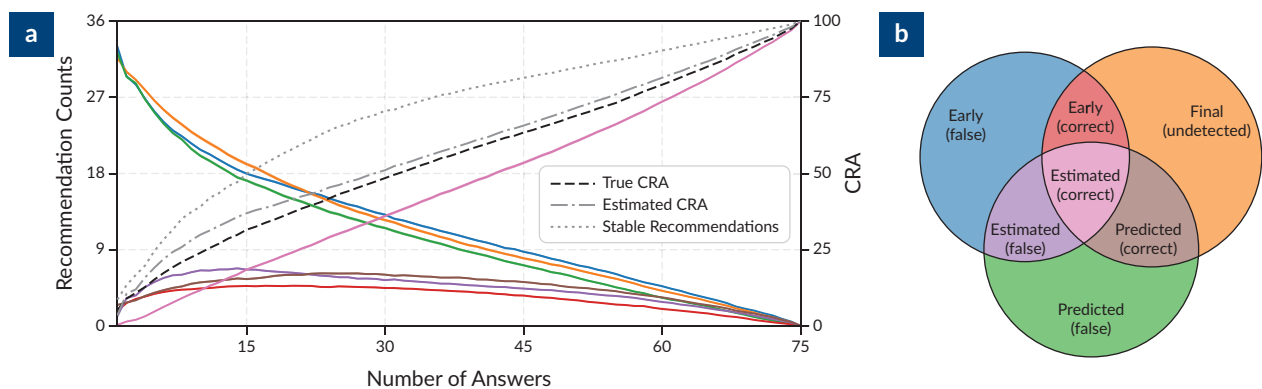


Figure 3. Venn diagram of the early, predicted, and final recommendations. Notes: (a) The blue, green, and orange lines show the number of candidates that are only in one of the three sets of recommendations; they decrease from 36 (when all sets are distinct) to 0 (when all sets are identical); the pink line shows the intersection of all three sets; the dotted line shows the fraction of stable recommendations in the estimated recommendations; (b) In the corresponding Venn diagram, the true positives (correct) and false positives (false) are colored according to the lines in the graph.

4.3.4. Optimal Thresholds for Sampled Recommendations

Figure 4 shows the number of true and false positives for each of the three approaches. While Early recommendations contain most true positives, they start overconfidently with around 35 false positives. In contrast, Estimated recommendations initially only select around three true and seven false positives. Sampled recommendations further reduce the number of false positives, such that, especially at the beginning of the questionnaire, there are very few false positives. Here, we identified a sample size of $N = 100$ and a threshold of $t = 0.7$ to be optimal (see Table 9 in the Supplementary File). As shown in Figure 4b, during the first five questions, the difference between true and false positives is minimal and has a low variance across voters. The bar at the bottom of the figure shows that the improvement of Sampled is statistically significant for the first 25 questions ($p < 0.001$). Then, the effect disappears, and the performance of both Sampled and Estimated equalizes.

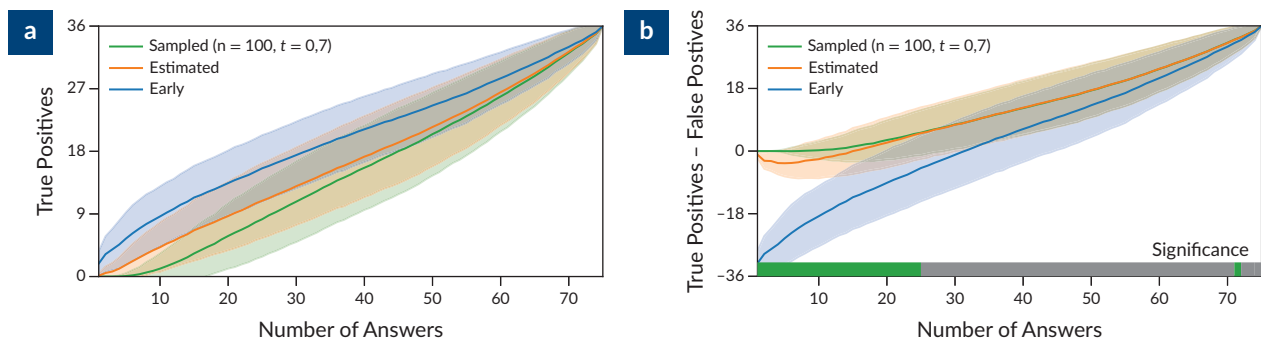


Figure 4. The precision of the stable recommendations for three different selection algorithms. Notes: (a) The y-axis shows the number of true positives per question across all voters (mean and standard deviation); the Early recommendations constantly have a higher score than the Sampled and Estimated ones; (b) The y-axis shows the difference between true and false positives per question across voters (mean and standard deviation); the Early recommendations start with many false positives; the horizontal bar shows where the difference between the Estimated (orange) and Sampled (green) recommendations is statistically significant.

5. Evaluating the Display of Recommendation Certainty

To evaluate the effects of displaying the estimated CRA and stable recommendations to users, we conduct a controlled user experiment on a self-developed VAA platform. Here, we implement the best-performing estimation algorithms and three different user interfaces within the platform, communicating the recommendation certainty to the experiment participants. This section presents the interfaces, the experimental setup, and the results of the user experiment.

5.1. Platform Design

To run the experiment, we developed a web-based platform specifically designed to run controlled experiments on user behavior in VAAs. It consists of two main components: the survey page and the candidate recommendation page. On the survey page, users react sequentially to the Smartvote questions. These reactions are on a 5-point Likert scale, including a neutral option. On the candidate recommendation page, users see the top matches with their party affiliation, similarity score, and position in the political map (see Figure 15 in the Supplementary File). Users can select which matches seem relevant to them. For simplicity, the number of candidates on the recommendation page is always limited to a relatively small set of $k = 10$ candidates. This way, users who are unfamiliar with VAAs are not overwhelmed by a large selection. A live online demo of the platform is accessible here: <https://aqvaa.ifi.uzh.ch/survey/3xqPQz>

We extend the survey page of the VAA platform with three semantic progress indicators: The first interface displays the estimated CRA with a progress bar below each question (see Figure 5). In contrast to a standard progress bar, this estimated CRA adapts to the users' answers and may increase or decrease after each question. The second and third interfaces display a set of stable recommendations below each question (see Figure 6). Depending on the estimation algorithm (Estimated and Sampled), the interface is slightly different. As shown in Figure 6a, the Sampled recommendations are shown together with their probability given by percentages. Only for the candidates whose frequency is above the threshold t are party and name displayed in a blue box. For candidates with a lower frequency, a question mark is displayed in a yellow box.

This is designed to spark curiosity while keeping the number of false positives minimal. As shown in Figure 6b, the third interface merely shows the Estimated recommendations in blue boxes without indicating their stability and corresponding percentage. Both these interfaces show a grey box for the fields that cannot yet be filled with stable recommendations.

Note that for all interfaces, we add a textual explanation that reads: “As you answer more questions, your recommendations become more stable.” This should encourage users to answer more questions, increasing the recommendation certainty. In addition, we disable the standard progress bar that typically shows the number of answered questions at the top. This way, we avoid people feeling urged to complete all 75 questions.

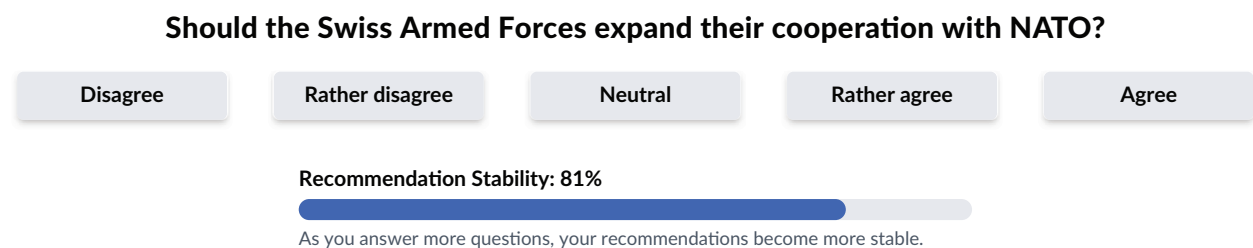


Figure 5. Interface of the estimated CRA in the VAA platform. Notes: At each question, the estimated CRA is displayed in the progress bar; depending on the response pattern, the certainty can increase or decrease with each answer; the question and response buttons are the same for each different interface.

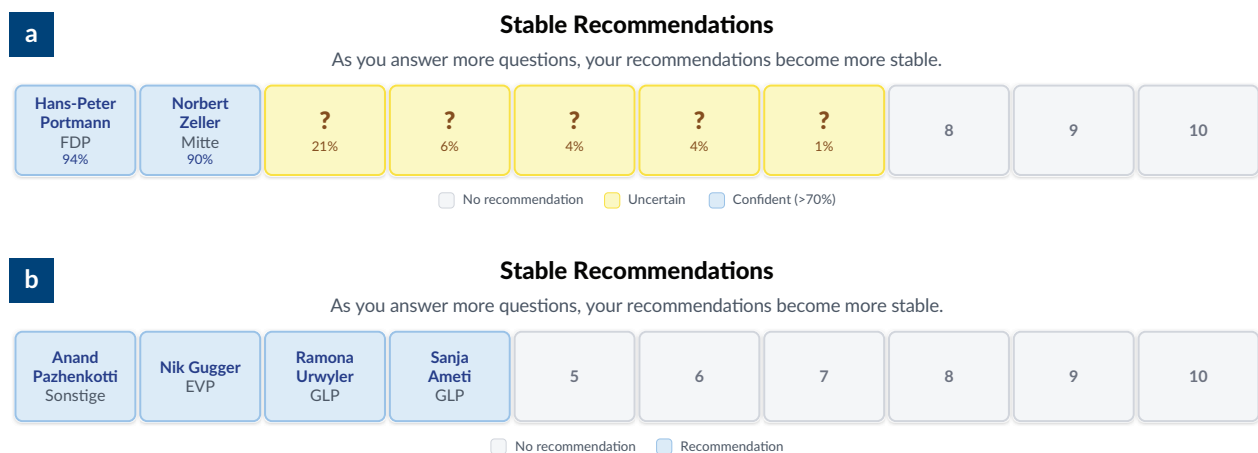


Figure 6. Interfaces of the stable recommendations in the VAA platform. Notes: (a) For the sampled recommendations, the algorithm computes the probability of each candidate being among the final recommendations; the name and party are displayed if the percentage is above a certain threshold; if it is below the threshold, just the probability is displayed; for the remaining recommendations, a gray box is shown; (b) For the estimated recommendations, only the overlap of early and predicted recommendations is highlighted in blue.

5.2. Experimental Setup

The following user experiment was reviewed and approved by the human subjects committee of our faculty. As shown in Figure 7, users start with a pre-survey for screening purposes. These questions, listed in Table 5 in the Supplementary File, collect demographic variables. Then, users are forwarded to our platform, where they answer the VAA questionnaire with the condition-specific interface. After answering a minimum of

15 questions, a button “Show recommendations” appears. By clicking this button, users can, from then on, exit the questionnaire to view their candidate recommendation page. A pop-up explains to the users that they can do so whenever they feel that they have answered enough questions. On the recommendation page, users then rate their recommendations for the first time (Q_1). After that, they must answer at least another five VAA questions before they can click “Show recommendations” for a second time (Q_2). This time, proceeding to the recommendation page terminates the questionnaire. Again, we explicitly explain this procedure with a pop-up. After submitting the second rating, users are forwarded to a post-survey, which collects information about their perception of the interface. The post-survey questions are listed in Tables 6 and 7 in the Supplementary File.

Users are randomly assigned to one of five conditions: With the first and second conditions, we address RQ2, exploring whether the displayed CRA estimate influences the point of dropout (Q_2). Both these conditions share the same interface with the progress bar displaying the recommendation certainty as shown in Figure 5. In one condition, however, users see the true estimated CRA (TrueCRA) while in the other, they see an artificially fast-growing estimated CRA (FastCRA). This way, we can evaluate the influence of the displayed value. With the third and fourth conditions, we address RQ3, investigating whether users perceive simpler interfaces as easier to understand than more accurate ones. Here, we display the stable recommendations to the users with the interfaces shown in Figure 6. In one condition, users see the estimated recommendations without percentages (Estimated), while in the other, users see the sampled recommendations with displayed probabilities for each candidate (Sampled). We set the threshold for these probabilities to $t = 0.7$ as preliminary results showed that this threshold maximizes the difference between true and false positives. Lastly, in the control group, users do not see any information about the recommendation certainty (Control).

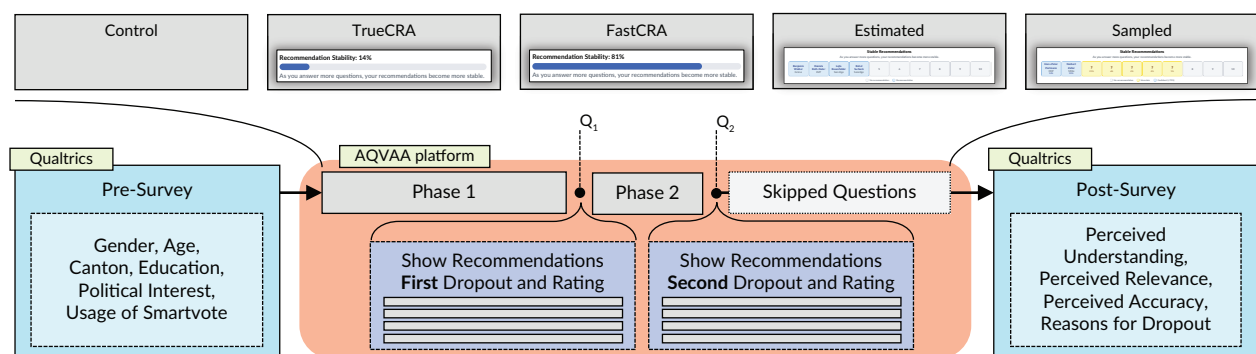


Figure 7. Flowchart of the user experiment. Notes: Participants start with a pre-survey on Qualtrics, where they are screened for demographic variables; then, they are assigned to one of 5 groups—Control, TrueCRA, FastCRA, Estimated, or Sampled—and forwarded to the VAA; on the VAA, users answer a minimum of 15 questions; then they can freely decide to see their recommendations (Q_1); after rating the recommendations, users proceed to answer more questions; after a minimum of 5 additional questions, they can again freely terminate the questionnaire to see their updated recommendations (Q_2); after rating them, users proceed to the post-survey, evaluating the interface and providing reasons to skip the remaining questions.

5.3. Results of the User Experiment

We recruited 130 participants from the Canton of Zurich through the market research company Bilendi. All participants (53 men and 77 women) completed the survey between the 24th and 30th of July 2025. The median survey duration was 14:12 minutes, where participants from the Estimated and Sampled condition took significantly longer than the control group ($p = 0.020$). We excluded from the analysis 27 users for one of the following reasons: either they did not understand the instructions (they stated that they completed all 75 questions to get the full financial incentive or were not aware that early dropout was possible), or they had an unreasonable duration of the survey (faster than 5 min or slower than 120 min). Despite random assignment to the experiment conditions, users in one condition (Estimated) report a significantly higher political interest than the Control group ($p = 0.011$). In our statistical analysis, we therefore control for political interest using an OLS model and an ANCOVA with the corresponding pairwise Tukey's HSD test.

5.3.1. Answered Questions Before Dropping Out

On average, Smartvote users answer 57 questions before proceeding to their candidate recommendations. In the Control group of our experiment, users answered on average 56 questions before dropping out. Across conditions, this number differs between participants: 88 users (67.7%) filled out all 75 questions, while 6 (4.6%) quit the questionnaire directly after the minimum of 20 questions. The rest was evenly distributed in between (see Figure 17 in the Supplementary File). We investigate whether the displayed recommendation certainty influenced participants' decision to finish the questionnaire. As Table 2 shows, Estimated has a significantly later dropout than the Control group ($p = 0.005$). Furthermore, TrueCRA, Estimated, and Sampled have a significantly later dropout than FastCRA ($p < 0.001$).

5.3.2. Perceived Influence of the Display

Contrary to the significantly different number of answered questions, users did not perceive that having the certainty displayed influenced their decision to click "Show recommendations." As shown in Table 2, most users reported that the certainty display had a neutral influence on them (with a slight negative tendency ranging from -0.7 to -0.2). However, the displayed recommendation certainty at dropout was lower for FastCRA than for the other three conditions. These differences, however, are not statistically significant.

Participants selected different reasons in the post-survey to explain their dropping out (a list of the pre-defined options is given in Table 7 in the Supplementary File). Users who quit early were mainly "satisfied with the certainty" (25%) or "looked at the certainty" (25%). Others saw "no incentive to continue" (25%) or gave other reasons (25%). In the open-ended feedback, one user wrote that "in hindsight, [they] would have done the full questionnaire." Those users who completed the whole questionnaire mainly did so "out of curiosity" (41%). Some wanted to have the recommendations "as precise as possible" (28%) and "looked at the certainty" (10%). In the open-ended feedback, one user reported that they "did not want to be influenced by any information." Another wrote: "The more information is available, the more precise (and therefore relevant) are the recommendations."

Table 2. User evaluation of the interfaces in the VAA platform.

Condition	Dropout	Final CRA	Displayed	Understanding	Influence	Relevance	Accuracy	Usage
Control (21)	56 ± 5	75 ± 7						0.8 ± 0.4
FastCRA (17)	49 ± 5	58 ± 8	77 ± 6	0.6 ± 0.5	−0.4 ± 0.5	0.8 ± 0.4	0.6 ± 0.4	1.2 ± 0.4
TrueCRA (22)	63 ± 4	86 ± 5	86 ± 5	0.9 ± 0.4	−0.2 ± 0.4	0.7 ± 0.4	0.2 ± 0.4	0.6 ± 0.4
Estimated (21)	68 ± 3	90 ± 4		1.5 ± 0.3	−0.7 ± 0.4	0.9 ± 0.4	0.5 ± 0.3	0.9 ± 0.4
Sampled (20)	65 ± 4	86 ± 6		0.3 ± 0.4	−0.5 ± 0.4	1.3 ± 0.2	0.6 ± 0.3	1.1 ± 0.3

Notes: The table shows means and their standard errors; the dropout refers to the total number of answered questions; only in FastCRA did the displayed CRA estimate differ from the estimated CRA; the remaining columns are results from the post-survey; values in bold indicate a significantly higher result than that of another condition, and values in italic a significantly lower result than that of another condition.

5.3.3. Understanding of the User Interface

Users reported understanding the Estimated interface significantly more than Sampled ($p < 0.001$) and FastCRA ($p = 0.018$). Perceived relevance and accuracy of the interfaces were not significantly different. However, there is a tendency that Sampled were found to be more relevant than all other conditions, including Estimated. Users in the FastCRA reported the highest likelihood of recommending the website to a friend or family member. Their evaluation was closely followed by Sampled.

Users in the Control group did not see the interfaces during the questionnaire. In the post-survey, we showed them screenshots of all interfaces and asked users to rate them hypothetically. Users reported that they understood the interface with the certainty bar better than the interface with stable recommendations; they also found it more interesting and more accurate. However, the number of participants in the control group (21) did not allow for statistically significant results in a paired t -test ($p = 0.201$, $p = 0.108$, and $p = 0.367$).

5.3.4. General Thoughts of Participants

The users in our experiment were not typical VAA users; 54 (52%) had never consulted a VAA before. As one user put it: “It would be great if such a tool were publicly available during cantonal elections or similar.” Yet, users’ overall experience of our VAA platform was positive. As shown in Table 8 in the Supplementary File, this was explicitly expressed by 25 users in the open-ended feedback. For example, one user wrote that they “found it interesting to get recommendations based on [their] answers.” Twelve users mentioned that they had learnt something using the VAA. One noted that they were “surprised, as a not very politically interested person, what kind of candidates were recommended.” Seven users commented on the displayed recommendation certainty. Among those, four users critiqued the explanations of the certainty display. One said that while they “found the questionnaire good, the certainty display was slightly confusing at first.” They thought that the certainty bar displayed the system’s evaluation of their responses. Three other users highlighted the interface positively. One wrote: “Since there was no (normal) progress bar, I focused on the certainty bar and was happy.”

6. Discussion

Overall, we found that users in our experiment appreciated the interfaces similarly and reported a high likelihood of recommending the platform. However, there are some differences across users and conditions that warrant further attention. We first address the results of the simulation, focusing on the accuracy of the algorithms and their variance across user types. We then address the results of the user experiment, focusing on the discrepancy between users' perception and their observed behavior on the platform.

6.1. Posterior-Informed Samples Predict Stable Recommendations

Our simulation results provided clear evidence that the recommendation certainty can be accurately estimated, addressing RQ1. The posterior-based samples yielded estimates that were, with an error of 7.95%, significantly more accurate than those of the heuristic baseline. Moreover, we found that the number of samples significantly affected the error and the fluctuation of the estimated CRA. In the trade-off between duration and performance, we therefore argue that the number of samples be set to $M = 100$ with a threshold of $t = 0.7$, which seems to be a point of diminishing returns.

The precision of the algorithms, however, was not noticed by all participants. Across conditions, users reported a similar perceived accuracy—even in FastCRA, which displayed an artificially high CRA estimate (see Figure 2b). Our interpretation is that users might have a vague view of “good” recommendations and therefore do not perceive false positives or overconfident CRA estimates as inaccurate.

6.2. Precision Differs Across User Types

The estimated CRA, as a measure of recommendation certainty, depends on users' individual response patterns. For example, Figure 8 shows the progression of the true and estimated certainty for user 3613 throughout the questionnaire. Despite sudden fluctuations (e.g., after five questions), the estimated CRA accurately follows the trend of the true CRA. However, there might be a larger error for other users (see

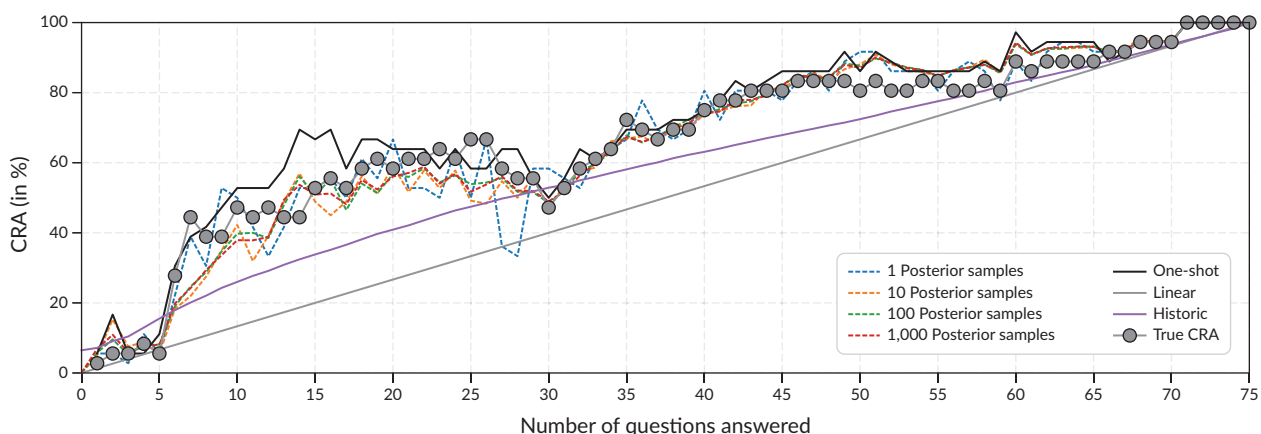


Figure 8. Progression of the true CRA compared to different estimation algorithms. Notes: The gray dots show the true CRA of user 3613; after questions 6 and 7, the CRA suddenly jumps to 42%; then, the increase in CRA is moderate until it drops again after 30 questions; for the remaining questions, the CRA increases almost monotonically until it reaches 100% after all questions; the colored lines indicate the estimated CRA for the different algorithms.

Figure 18 in the Supplementary File). We identified three possible explanations for this behavior: First, these users tend to answer the questionnaire with an atypical response pattern; second, they often answer the questions with relatively neutral stances; third, the number of candidates in their proximity is very high. If a user, for example, holds political views that few candidates share, these candidates will quickly emerge as the corresponding recommendations. In contrast, if many similar candidates could equally well be recommended, sampling the remaining answers affects the choice among these candidates, thereby reducing the certainty of the recommendations. This, however, is not due to the quality of the method, but is instead an artifact of the political landscape itself, which lacks a uniform density of candidates.

6.3. Curiosity Keeps Users in the Questionnaire

The user experiment demonstrated a significant effect of the displayed recommendation certainty on user behaviour, addressing RQ2. Even though users reported that the display did not influence the number of answered questions, they still dropped out significantly earlier in FastCRA than in TrueCRA. In addition to this result, we also found that users in Estimated and Sampled spent more time on the survey and answered more questions than users in the conditions with progress bars. We take this as an indication that stable recommendations sparked users' curiosity. Displaying intermediate recommendations may be a cognitive anchor for users to reflect more deeply on their responses. It remains to be investigated, however, whether such a display would influence their answers in a specific direction through some form of confirmation bias or reassurance. This concept is discussed in the literature as performative prediction and calls for future research in the context of political questionnaires (Mendler-Dünner et al., 2022; Perdomo et al., 2020).

6.4. Users Prefer Simple Over Accurate Displays

We furthermore found that users perceived their understanding of each interface differently, as inquired by RQ3. The display of Estimated recommendations, which just shows candidates' names and parties, was perceived as most understandable. The display of Sampled recommendations, which also shows individual candidates' certainties as percentages, was perceived to be the least understandable. As the percentages are the main difference between the interfaces, this result indicates that such additional information was difficult for users to interpret. However, while the Estimated interface is simpler, it is also less accurate, since the simulation results revealed more false positives in that condition. Interestingly, users across all conditions rated their interest in using the platform in a similar way, indicating that their understanding of the interface did not impact their experience.

6.5. Generalizability Beyond Candidate-Based VAAs

Lastly, we address the generalizability of the study. While we focused our analysis on candidate-based VAAs, the method can also be applied to party-based VAAs—or any application that recommends a target variable based on interactions in a sequential questionnaire. However, we identify two adjustments for party-based VAAs: First, the CRA needs to be changed to another metric, such as Spearman's rank correlation, as there are no "sets" of candidates but ranked lists with recommended parties or single recommendations. Moreover, the smaller size of the training data might limit the quality of the model's predictions. This can, however, be addressed by using historic voter-question interactions or by generating training data with large language models (Bachmann et al., 2025).

7. Limitations

Our study has several limitations. We divide them into limitations of the simulation framework and limitations of the user experiment.

7.1. Simulation Framework

The most apparent limitation of our method is that we measure the certainty as the overlap of early and final recommendations, which might not reflect nuances of the recommendation quality. First, this metric does not include the ranking of candidates, which might be important to some users. Second, we ultimately do not know what a “good recommendation” is, i.e., it might be that the final recommendations are not perceived as more relevant than early ones. However, we believe this approach is the best approximation of the ground truth, while acknowledging that users might perceive it differently.

We used a simplified simulation of the VAA to evaluate the CRA estimation algorithms. For example, our analysis focused on a specific size of recommendation set ($k = 36$ and $k = 10$). However, in some cantons, different numbers of candidates are recommended. While the overall trend of the results is the same for $k = 10$, we note that the CRA estimates are less accurate and fluctuate more for smaller k . Moreover, our choice of statistical model might not be optimal. While the combination of PCA and LR yields promising results, it entirely ignores the fact that users often prioritize certain questions (weighing) and that the ordinal responses in the original data do not necessarily map to the continuous predictions of the model. Lastly, a general limitation of our method is its dependency on training data. These data are specific to each election and questionnaire. Even if training data naturally arise in candidate-based VAA through the responses of the eligible candidates, the model's performance could be affected if there is a different response pattern in the candidates' and users' answers. This could decrease the predictive accuracy for voters.

7.2. User Experiment

There are also limitations to our user experiment. First, we designed the interfaces focusing more on technical accuracy than user friendliness. While we tried to provide an accessible platform, some participants got stuck on the candidate recommendation page and did not know how to proceed to the post-survey. Moreover, the participants were recruited through a market research company and received a financial incentive to complete the questionnaire, rather than for personal interest. We therefore had to exclude a substantial number of participants from the analysis, resulting in a relatively modest sample size, which is not representative of typical Smartvote users (who are younger and more liberal) and does not represent Zurich's political and demographic population (our participants are older and slightly more conservative). However, we argue that this exploratory research is valuable because it does not survey the typical Smartvote user base, who already use the tool.

A different experimental setup—such as a 2×2 study design—would have improved the evaluation of the interfaces. Instead of only showing an artificially fast-growing CRA in the certainty bar interface (TrueCRA vs. FastCRA), we could have also displayed an artificially large number of stable recommendations in the other interfaces. Such a setup would have allowed the analysis of cross-effects of both treatments. For a more complete understanding of the results, it would have been beneficial to collect more data on *why* people decided to drop out and see their recommendations. Instead, we collected reasons through a multiple-choice

post-survey. Lastly, we did not measure users' *actual* understanding of the interfaces; we just asked for their *perceived* understanding. Despite these limitations, the results we do have provide exploratory insights into the implementation of semantic progress indicators in VAAs and can be used to inform future research.

8. Conclusion

In this study, we explored the integration of a semantic progress indicator into a candidate-based VAA. We developed a method to estimate the certainty of early recommendations and implemented this algorithm with three different interfaces in our VAA. Through a controlled user experiment, we found that this additional information can spark curiosity and lead to prolonged user engagement.

The core idea of the study—using different types of predictions to estimate their accuracy—applies to many fields outside VAAs. While previous research has shown that the predicted recommendations are slightly more accurate than the early recommendations (Bachmann et al., 2024), the unexpected finding was that their overlap indicates the quality of both. In other words, the predicted recommendations uncover different candidates from the early recommendations if and only if both sets are far from the final set. This geometric phenomenon makes the topic interesting from a mathematical perspective.

From a political perspective, it was interesting that users cared more about the interface in which the estimate was displayed than about its accuracy. This finding suggests that, for interface design, simplicity is often preferable, even at the expense of technical precision. As a promising trade-off, the preview of estimated recommendations emerged—an interface that was well understood, while having satisfactory properties regarding true and false positives.

Thinking ahead, this research on estimating the CRA not only informs the feature design in candidate-based VAAs but can also help to personalize adaptive strategies through the model's predictions of users' individual progression of the CRA. The different types of CRA curves (smooth, surprising, and unpredictable) could be formalized into a typology of user profiles and contribute to political research on citizens' voting patterns in contrast to those of political elites.

Acknowledgments

First, we would like to thank Politools for providing the Smartvote data and offering invaluable advice on the study design. Moreover, we would like to thank all students who helped implement the VAA and conduct the experiments. Lastly, we thank the reviewers, editors, and everyone who provided feedback on the early drafts of this work, thereby helping to improve it.

Funding

We gratefully acknowledge the support of the Swiss National Science Foundation (SNSF) under grant ID CRSII5-205975, which provided the primary funding for our research. Publication of this article in open access was made possible through the institutional membership agreement between the University of Zurich and Cogitatio Press.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

All data and code necessary to reproduce the results are publicly available in this GitHub repository: <https://github.com/fsvbach/recommendations-pag-paper>. Please contact the authors directly with potential questions.

LLMs Disclosure

Generative AI was used to improve and document our research code during the analysis, and to revise elements of the text for grammar and clarity.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Agathokleous, M., & Tsapatsoulis, N. (2016). Applying hidden Markov models to voting advice applications. *EPJ Data Science*, 5(1), Article 34. <https://doi.org/10.1140/epjds/s13688-016-0095-z>
- Bachmann, F., Sarasua, C., & Bernstein, A. (2024). Fast and adaptive questionnaires for voting advice applications. In A. Bifet, T. Krilavičius, I. Miliou, & S. Nowaczyk (Eds.), *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track* (pp. 365–380). Springer Nature. https://doi.org/10.1007/978-3-031-70381-2_23
- Bachmann, F., van Der Weijden, D., Heitz, L., Sarasua, C., & Bernstein, A. (2025). Adaptive political surveys and GPT-4: Tackling the cold start problem with simulated user interactions. *PLoS One*, 20(5), Article e0322690. <https://doi.org/10.1371/journal.pone.0322690>
- Berdoz, F., Brunner, D., Vonlanthen, Y., & Wattenhofer, R. (2025). Recommender systems for democracy: Toward adversarial robustness in voting advice applications. In James Kwok (Eds.), *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence* (pp. 9564–9572). International Joint Conferences on Artificial Intelligence Organization. <https://doi.org/10.24963/ijcai.2025/1063>
- 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>
- Buryakov, D., Kovacs, M., Serdült, U., & Kryssanov, V. (2024). Enhancing the design of voting advice applications with BERT language model. *Frontiers in Artificial Intelligence*, 7, Article 1343214. <https://doi.org/10.3389/frai.2024.1343214>
- Buuren, S. V. (2012). *Flexible imputation of missing data*. CRC Press.
- Card, S. K., Robertson, G. G., & Mackinlay, J. D. (1991). The information visualizer, an information workspace. In S. P. Robertson, G. M. Olson, & J. S. Olson (Eds.), *CHI '91: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 181–186). Association for Computing Machinery. <https://doi.org/10.1145/108844.108874>
- Clinton, J., Jackman, S., & Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2), 355–370. <https://doi.org/10.1017/S0003055404001194>
- Elahi, M., Ricci, F., & Rubens, N. (2016). A survey of active learning in collaborative filtering recommender systems. *Computer Science Review*, 20, 29–50. <https://doi.org/10.1016/j.cosrev.2016.05.002>
- Garzia, D., & Marschall, S. (2019). Voting advice applications. In *Oxford research encyclopedia of politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.620>
- 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>

- 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. (2016). Dynamic scale validation reloaded: Assessing the psychometric properties of latent measures of ideology in VAA spatial maps. *Quality & Quantity*, 50(3), 981–1007. <https://doi.org/10.1007/s11135-015-0186-0>
- Grung, B., & Manne, R. (1998). Missing values in principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 42(1/2), 125–139. [https://doi.org/10.1016/s0169-7439\(98\)00031-8](https://doi.org/10.1016/s0169-7439(98)00031-8)
- Kamoen, N., & Liebrecht, C. (2022). I need a CAVAA: How conversational agent voting advice applications (CAVAAs) affect users' political knowledge and tool experience. *Frontiers in Artificial Intelligence*, 5, Article 835505. <https://doi.org/10.3389/frai.2022.835505>
- 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>
- Kay, M., Kola, T., Hullman, J. R., & Munson, S. A. (2016). When (ish) is my bus? User-centered visualizations of uncertainty in everyday, mobile predictive systems. In J. Kaye, A. Druin, C. Lampe, D. Morris, & J. P. Hourcade (Eds.), *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5092–5103). Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858558>
- Ladner, A., Fivaz, J., & Pianzola, J. (2012). Voting advice applications and party choice: Evidence from SmartVote users in Switzerland. *International Journal of Electronic Governance*, 5(3/4), 367–387.
- Law, E., Yin, M., Goh, J., Chen, K., Terry, M., & Gajos, K. Z. (2016). Curiosity killed the cat, but makes crowdwork better. In J. Kaye, A. Druin, C. Lampe, D. Morris, & J. P. Hourcade (Eds.), *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4098–4110). Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858144>
- 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>
- Lü, L., Medo, M., Yeung, C. H., Zhang, Y.-C., Zhang, Z.-K., & Zhou, T. (2012). Recommender systems. *Physics Reports*, 519(1), 1–49. <https://doi.org/10.1016/j.physrep.2012.02.006>
- McCoy, J. T., Kroon, S., & Auret, L. (2018). Variational autoencoders for missing data imputation with application to a simulated milling circuit. *IFAC-PapersOnLine*, 51(21), 141–146. <https://doi.org/10.1016/j.ifacol.2018.09.406>
- Mendler-Dünner, C., Ding, F., & Wang, Y. (2022). Anticipating performativity by predicting from predictions. *Advances in Neural Information Processing Systems*, 35, 31171–31185. <https://dl.acm.org/doi/10.5555/3600270.3602530>
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706. <https://doi.org/10.1080/10584609.2020.1843572>
- Perdomo, J. C., Zrnic, T., Mendler-Dünner, C., & Hardt, M. (2020). Performative prediction. In H. Daumé & A. Singh (Eds.), *ICML '20: Proceedings of the 37th International Conference on Machine Learning* (Vol. 119, pp. 7599–7609). PMLR. <https://proceedings.mlr.press/v119/perdomo20a.html>
- Pianzola, J., Trechsel, A. H., Vassil, K., Schwerdt, G., & Alvarez, R. M. (2019). The impact of personalized information on vote intention: Evidence from a randomized field experiment. *The Journal of Politics*, 81(3), 833–847. <https://doi.org/10.1086/702946>
- Politools. (2023). *Daten zu den Nationalrats- und Ständeratswahlen 2023 der Online-Wahlhilfe Smartvote* [Data set]. Smartvote. <https://www.smartvote.ch>

- Poole, K. T., & Rosenthal, H. (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science*, 29(2), 357–384. <https://doi.org/10.2307/2111172>
- Potthoff, R. (2018). Estimating ideal points from roll-call data: Explore principal components analysis, especially for more than one dimension? *Social Sciences*, 7(2), Article 12. <https://doi.org/10.3390/socsci7010012>
- Romero Moreno, G., Padilla, J., & Chueca, E. (2022). Learning VAA: A new method for matching users to parties in voting advice applications. *Journal of Elections, Public Opinion and Parties*, 32(2), 339–357. <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 & Internet*, 8(4), 431–456. <https://doi.org/10.1002/poi3.139>
- Rubens, N., Elahi, M., Sugiyama, M., & Kaplan, D. (2015). Active learning in recommender systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender systems handbook* (pp. 809–846). Springer. https://doi.org/10.1007/978-1-4899-7637-6_24
- Schwarz, D., Schädel, L., & Ladner, A. (2010). Pre-election positions and voting behaviour in parliament: Consistency among Swiss MPs. *Swiss Political Science Review*, 16(3), 533–564. <https://doi.org/10.1002/j.1662-6370.2010.tb00440.x>
- Swiss Federal Statistical Office. (2023). *Federal elections*. <https://www.wahlen.admin.ch/en/zh>
- Tromborg, M. W., & Albertsen, A. (2023). Candidates, voters, and voting advice applications. *European Political Science Review*, 15(4), 582–599. <https://doi.org/10.1017/S1755773923000103>
- Urbistat. (2025). *Age classes by gender region Zurich. Maps, analysis and statistics about the resident population*. <https://ugeo.urbistat.com/AdminStat/en/ch/demografia/eta/zurich/1/2>

About the Authors



Fynn Bachmann is a PhD candidate in computer science at the University of Zurich. Previously, he obtained a master's degree in machine learning from the University of Tübingen and a bachelor's degree in physics from the University of Heidelberg. More information can be found on his personal website: <https://fynnbachmann.com>



Daan van der Weijden is a PhD candidate in computer science at the University of Zurich. He obtained a master's degree in artificial intelligence and a double bachelor's degree in artificial intelligence and linguistics, both from the University of Utrecht. More information can be found on his personal website: <https://daanvdweijden.com>



Cristina Sarasua is a senior researcher in computer science at the University of Zurich. Her research lies at the intersection of human-centered computing, data-driven technology, and civic participation. Prior to joining the University of Zurich, Cristina worked at the University of Koblenz-Landau, where she received her PhD in computer science.



Abraham Bernstein is a full professor of informatics at the University of Zurich. His current research focuses on various aspects of artificial intelligence and human-computer interaction with practical applications in media, medicine, and political science. His work is based on both social science and technical foundations. Mr. Bernstein is also a founding director of the University of Zurich's initiative for the digital transformation (called DSI) involving more than 1,400 researchers from all disciplines. For more information, see his research group's webpage at <https://ddis.ch>

In Pursuit of Informed Voters: Three Experimental Studies on Enhanced Voting Advice Applications

Naomi Kamoen ¹ , Christine Liebrecht ¹ , and Rieke van Lieshout ² 

¹ Department of Communication and Cognition, Tilburg University, The Netherlands

² Faculty of Arts, Radboud University, The Netherlands

Correspondence: Naomi Kamoen (n.kamoen@tilburguniversity.edu)

Submitted: 29 August 2025 **Accepted:** 27 November 2025 **Published:** 11 February 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Voters frequently struggle to understand political attitude statements in voting advice applications (VAAs) and often invest limited effort in resolving these difficulties. Conversational agent VAAs (CAVAAs) aim to reduce the cognitive effort involved in searching for relevant information by integrating chatbots that can provide contextual support. This article presents findings from three studies comparing CAVAAs to standard VAAs without additional information (Studies 1 and 2) and to a VAA with static clickable information (VAA+, Study 3). Study 1 ($N = 93$) was a laboratory experiment conducted during the 2023 Dutch parliamentary elections. University students were assigned to a standard VAA or a CAVAA. The chatbot in the CAVAA was used in approximately 45% of cases, with users showing a preference for opinion-based and status quo information. Compared to the VAA, CAVAAs reduced non-directional responses to the VAA statements and increased evaluations of the tool's usability. Moreover, users' perceived knowledge was higher, whereas no differences were found for factual knowledge and turnout intention. Study 2 ($N = 144$) largely replicated these results outside the lab with a more diverse sample and showed that the effects hold across different levels of political sophistication. Study 3 ($N = 159$), conducted during the 2024 European elections, compared a CAVAA to a VAA+. While VAA+ users requested information more frequently, both tools received similar evaluations, and this finding was again consistent across groups of different political sophistication. In the manuscript, we discuss the implications of these findings for theory and practice.

Keywords

conversational agents; political knowledge; political sophistication; voting advice applications

1. Introduction

The 2022 Dutch municipal elections saw a historically low turnout, with only 50.99% of eligible voters voting. In a letter to the House of Representatives, the Minister of Interior and Kingdom Relations identified a lack of political knowledge as one of the primary reasons for this outcome (Bruins Slot, 2022). This conclusion aligns with Delli Carpini and Keeter's (1996) seminal work, which argues that a lack of political knowledge is among the most important reasons to abstain from voting.

Low levels of political knowledge and voter turnout are not exclusive to Dutch municipal elections; these issues are also prevalent in other countries and types of elections (e.g., Fivaz & Nadig, 2010). In today's increasingly complex political landscapes—particularly in multiparty systems like the Netherlands—voters often struggle to make informed decisions. One reason for this is the increasing number of political parties competing in elections. To illustrate this, the 2021 Dutch parliamentary elections featured 37 parties, 20 of which were new (Sipma et al., 2021). Other contributing factors include a growing sense of distance between citizens and politics, declining motivation to seek out information (see the literature review in Kamoen et al., 2015), and the rising influence of misinformation complicating voters' ability to assess the reliability of political content (Mason et al., 2018).

Parallel to these developments runs the steep rise of voting advice applications (VAAs) as a source of political information (Garzia & Marschall, 2012). In a VAA, voters express their opinions on approximately 30 political statements. The tool then compares the users' stances to those of the political parties or candidates, depending on what is relevant for the elections at stake, and calculates the strength of the match with each. The goal of a VAA, however, is not to calculate matches and make suggestions about whom voters could vote for. Instead, a VAA is an information resource aiming to provide reliable information about the political parties and their stances (De Graaf, 2010).

Empirical studies suggest that VAAs indeed increase users' (perceived) political knowledge and likelihood to vote (Munzert & Ramirez-Ruiz, 2021), especially among theoretically educated users (Heinsohn et al., 2019). At the same time, however, users have been found to encounter comprehension difficulties while interacting with these tools (Kamoen & Holleman, 2017). Although additional information to resolve these comprehension problems is available online, users rarely search for it and instead rely on assumptions to fill in the gaps of information.

To help VAA users obtain reliable political information at low cognitive costs, researchers have introduced a new type of VAA that includes a conversational agent (CA) in the form of a chatbot that is trained to respond to user queries related to political attitude statements in the VAA (e.g., Dieing, 2025; Hankel et al., 2024; Kamoen & Liebrecht, 2022; Van Veggel et al., 2025; Van Zanten & Boumans, 2023). Although CAVAAs, to the best of our knowledge, have not (yet) been launched in real-life elections, a first experimental comparison between a CAVAA and a VAA showed that the former yielded better user evaluations and higher political knowledge scores (Kamoen & Liebrecht, 2022). While these results are promising, the study had certain limitations, particularly in terms of ecological and construct validity, highlighting the need for further research.

Therefore, the present series of three studies further explores the usage of chatbots in the context of VAAs and builds on this earlier work by addressing its limitations and extending it in several ways. More specifically,

Study 1 compares a CAVAA to a VAA among theoretically-educated voters, using construct-valid experimental materials in the context of the 2023 Dutch parliamentary elections. Study 2 makes the same comparison, but among a diverse user audience. Finally, while Studies 1 and 2 compare a CAVAA to a VAA without supplementary information, Study 3 contrasts it with an enhanced VAA (VAA+) to understand if adding additional information via buttons in the standard web environment works as well as adding information via a chatbot. Across all studies, both tool evaluation measures (ease of use, usefulness, and playfulness) and political engagement measures (perceived and factual political knowledge, internal political efficacy, and turnout intention) are assessed as outcome measures. Additionally, the response behavior to VAA statements is included as an unobtrusive indicator of comprehension difficulties. As such, the aim is to advance theories of information processing in contexts where citizens invest minimal effort, while also providing insights for VAA practice regarding whether and how chatbots can be effectively integrated into this context.

2. Theoretical Framework

2.1. Comprehension Problems in VAAs

In a large-scale think-aloud study, Kamoen and Holleman (2017) found that VAA users encounter comprehension difficulties with roughly one in five statements. These comprehension problems ranged from difficulties with word meaning ("What is the VAT rate?"), to a lack of knowledge about the status quo ("How high is the VAT rate?"), to missing contextual arguments ("Why increasing the VAT rate?"). While VAAs aim to inform users (De Graaf, 2010), they do not always provide structured background information to resolve such problems. As a result, users must consult search engines or other external resources to fill these gaps. VAA users, however, have been found to display so-called satisficing behavior, opting to answer with limited effort, and thus limited information, rather than seeking clarification (Kamoen & Holleman, 2017).

The concept of satisficing stems from Krosnick's satisficing-optimizing model (Krosnick, 1991; Roberts et al., 2019), which explains variation in how respondents answer questions in surveys. While some aim to match their true opinion (optimizers), others opt for the quickest plausible answer (satisficers). Satisficing behavior is more likely when tasks are difficult, motivation is low, and when respondents have limited cognitive ability (Krosnick, 1991). While VAA users are motivated enough to enter a VAA survey voluntarily, these users, too, have been found to satisfice. In a VAA context, the satisficing behavior manifests in at least two ways. First, when facing comprehension problems, users tend to make assumptions rather than look up information (Kamoen & Holleman, 2017). Second, it amounts to the interchangeable usage of the "neutral" and "no opinion" response option to the VAA statements (Baka et al., 2012; Van Outersterp et al., 2016). This latter behavior is particularly problematic for those VAAs that assign different value to these responses in result calculations; for example, in the Dutch Stemwijzer, neutral answers are matched to parties, while no opinions are excluded from the calculation of the advice (De Graaf, 2010). As the voting advice has been found to affect the vote choice made (Wall et al., 2014), the interchangeable usage of these response options may lead to distorted matches and, ultimately, biased voting decisions.

2.2. Conversational Agent Voting Advice Applications

The comprehension challenges that VAA users face, combined with their limited motivation to consult external sources, provide two key reasons for integrating additional information directly into the VAA

interface. Researchers have therefore recently proposed integrating a CA in the form of a chatbot into a VAA, resulting in what is known as a Conversational Agent Voting Advice Application (CAVAA; e.g., Kamoen & Liebrecht, 2022).

Chatbots can take different forms, but most previous CAVAA studies (Hankel et al., 2024; Kamoen & Liebrecht, 2022; Van Veggel et al., 2025) have included buttons within the chatbot interface to access frequently asked questions. For example, for a statement like “The VAT rate on fruit should be increased,” users could click a button to learn what “VAT rate” means, another one to access information about the current rate, and a third one to learn about the arguments used in the political debate. In addition, users could type questions in an open chat. AI technology was then used to recognize these questions and return predetermined researcher-controlled responses (e.g., “Party X supports VAT reduction”). In case of unrecognized input, the chatbot would return an error message (e.g., “I do not understand your question”).

The theoretical advantage of such a chatbot, relative to a VAA without additional information, is that it mitigates users’ tendency to minimize effort. Since the information is embedded within the tool, users do not need to switch screens or search elsewhere. Moreover, the information is tailored to user needs: for example, those interested in the stances of Party A are not shown irrelevant information about Party B. Because of these presumed advantages, Kamoen and Liebrecht (2022) hypothesized that a CAVAA would be evaluated better than a VAA without information on tool evaluation measures (higher perceived usefulness, ease of use, and playfulness), as well as on political engagement measures (increased perceived and factual political knowledge, and turnout intention). In their Study 1, they indeed observed higher tool evaluation scores as well as higher perceived and factual political knowledge scores for the CAVAA. Counter to expectations, however, the CAVAA did not increase turnout intention.

While promising, a threat to the construct validity of that study was that the VAA and CAVAA had different layouts, which may have contributed to the observed differences. Furthermore, the study was conducted outside of an election context, whereas VAAs are typically used in the run-up to elections (Van de Pol et al., 2014). Therefore, the current Study 1 re-examines the hypotheses formulated by Kamoen and Liebrecht (2022), now using ecologically valid materials within the context of the 2023 Dutch parliamentary elections.

H1: Compared to users of a traditional VAA without information, CAVAA users will provide more positive tool evaluations.

H2: Compared to users of a traditional VAA without information, CAVAA users will be more politically engaged.

In addition to comparing VAAs to CAVAAs on these measures, the current research introduces the users’ response behavior to political attitude statements as an outcome measure. Based on previous studies (Baka et al., 2012; Van Outersterp et al., 2016), we expect that CAVAA users will provide neutral and no-opinion answers (jointly referred to as non-directional answers) less frequently to the political attitude statements, as these answers are both a sign of the comprehension difficulties and of the satisficing behavior CAVAAs aim to combat.

H3: Compared to users of a traditional VAA without information, CAVAA users will provide non-directional answers to VAA statements less frequently.

2.3. Political Sophistication

While Study 1 will test H1–H3 for theoretically educated voters, the goal of Study 2 is to compare the effect of CAVAA versus VAA for two types of users: those with higher and lower levels of political sophistication. Political sophistication is defined (e.g., Luskin, 1990) as a combination of cognitive ability (often proxied by education level), existing political knowledge (operationalized by a knowledge test), and motivation to acquire political information (typically operationalized as political interest). Distinguishing between users with higher and lower levels of sophistication is relevant because these groups are likely to differ in several respects, such as in the number and type of comprehension problems they encounter, their ability to recognize and articulate such problems (Elling et al., 2012), and their susceptibility to satisficing behavior (Krosnick, 1991). Moreover, real-world VAAs attract a broad user base (Van de Pol et al., 2014), making this distinction practically meaningful too.

In their study, Kamoen and Liebrecht (2022) also explored differences between users with higher and lower political sophistication. They did not formulate a hypothesis regarding this interaction, as theoretical arguments could be made in both directions: on the one hand, lower-sophisticated users might benefit more from CAVAA due to experiencing more comprehension difficulties; on the other hand, these users might struggle to recognize their information needs or to formulate appropriate questions for the chatbot, which could limit their ability to benefit from the tool. Results of their study found no interaction effect between user group and tool type on both tool evaluation measures and political engagement measures, suggesting that CAVAA may be equally effective for different types of users. However, the authors noted that their sample was skewed toward theoretically-educated participants (80%) and recommended replication with a more diverse population.

Therefore, Study 2 compares a CAVAA and a VAA among a heterogeneous population with more variation in educational level. Following earlier research, we do not formulate an hypothesis about the interaction of the effect of tool type and the level of political sophistication, but instead explore the relationship. Moreover, to better understand any possible effects of political sophistication, we will also explore any possible differences in the amount of information requested in the CAVAA between users with higher and lower levels of political sophistication.

2.4. Providing Information Within an Enhanced VAA+

While some VAA brands, such as the Dutch Kieskompas and the German Wahl-O-Mat, typically present only political attitude statements, other VAA brands offer additional information along with the political attitude statements in their VAAs. The Dutch VAA StemWijzer, for example, offers clickable explanations of key terms, a “Do you want to know more?” button that provides background information for each statement and a “party stances” section where users can view parties’ positions along with their justifications. Similar features are found in other European tools, such as Smartwielen in Luxembourg and Smartvote in Switzerland. This variation raises a question whether the CAVAA continues to provide added value when it is compared with a so-called VAA+, which offers political background information in a more conventional format through clickable buttons in the web environment.

This question is relevant for VAA design practice, but also from a theoretical perspective. Research on CAVAAAs can be linked to Krosnick's (1991) satisficing-optimizing model from survey research, as well as to the task-technology fit (TTF) theory from the field of information systems (Goodhue & Thompson, 1995). According to TTF, technology is most effective when its functionalities align with both the demands of the task and the characteristics of the user. In the context of VAAs, the task of forming an opinion on complex political statements can be cognitively demanding, especially for users with lower levels of political sophistication. While both CAVAAAs and VAA+s aim to support understanding by offering information, they differ fundamentally in structure and interactivity, so on the level of the "technology" in the TTF. VAA+s present predefined, static content through a familiar interface with clickable explanations. In contrast, CAVAAAs enable dynamic information retrieval, which may be more flexible but also more cognitively demanding. Study 3 explores how the format used to provide background information, chatbot (CAVAA) versus clickable web interface (VAA+), may affect user evaluation measures and political engagement measures, and whether this effect depends on users' political sophistication.

3. Method Study 1

3.1. Design

Study 1 was an experimental study conducted in the run-up to the 2023 Dutch parliamentary elections. Participants were randomly assigned to a traditional VAA without supplementary information or a CAVAA. In the original study set-up, there were two CAVAA presentation forms that we offered: either the chatbot was opened by default (proactive) or it was available through a mouse click (reactive). For consistency between the three studies, we only report on the comparison between the proactive CAVAA and the VAA; a comparison between all three conditions is reported in Kamoen et al. (2024). After using a VAA or a CAVAA, participants completed a survey that included both tool evaluation measures and political engagement measures. In addition, for each VAA statement, we recorded whether participants provided a directional or non-directional response to the VAA statements, and we logged which information was requested via the CA. The study was approved by the Research Ethics and Data Management Committee of the Tilburg School of Humanities and Digital Sciences at Tilburg University (reference number: REDC 2023.13a).

3.2. Materials

3.2.1. The VAA

The VAA was a copy of the Dutch StemWijzer. This copy contained 30 statements that VAA developer ProDemos created to represent the relevant issues in the Dutch political environment. In the tool, users are to indicate to what degree they agree with 30 political statements (e.g., "People aged 65 and older should be able to travel for free on trains") using a three-point scale (agree/neither agree nor disagree/disagree), or they may indicate that they want to skip the question. After having answered all statements, a "voting advice" is provided, depicted as a percentage of agreement with the parties running in the elections. The three most agreed-upon parties are highlighted at the top of the page.

The VAA differed from the original StemWijzer only at two points. First, the additional information about each VAA statement that was provided in the original was deleted from the VAA, as the goal was to make a comparison with an empty VAA. Second, the name, logo, and house style of the original StemWijzer were altered so that it was clear to participants that they were filling out an experimental VAA. ProDemos granted permission to use the copy in this altered form for the current study (Figure 1).

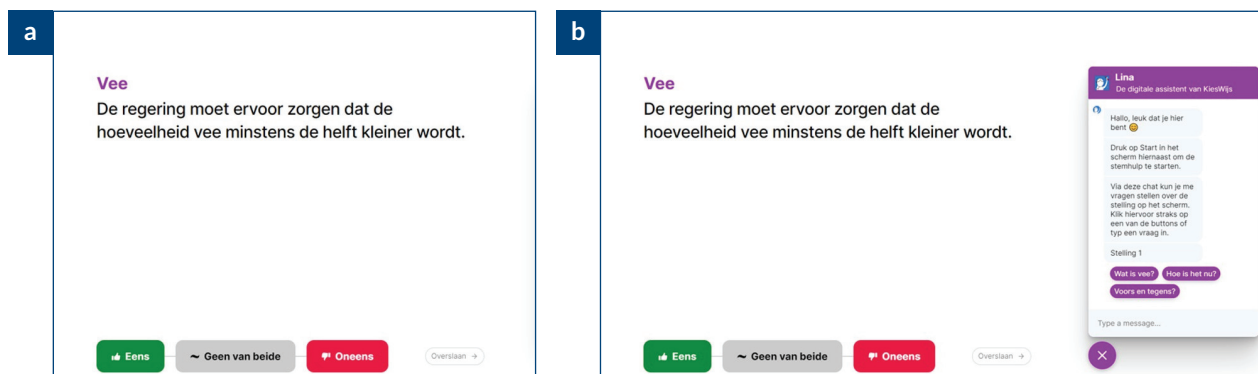


Figure 1. Screenshot of the VAA (a) and CAVAA (b). Notes: The header reads “Cattle” and the statement is “The government must ensure that livestock numbers are reduced by half.” In the CAVAA, the chatbot introduced itself and then provided three buttons and the open chat.

3.2.2. The CAVAA

In the CAVAA, a CA in the form of a chatbot was added to the VAA. The chatbot introduced itself on the first page of the tool and invited users to ask questions about the political attitude statements. For each statement, users could click on one of three suggestion buttons to access additional information (Figure 1b). The first button read “What is X?” where “X” was replaced by a key term (such as “Cattle”). The second button read “What is the situation now?” and clicking would lead to information about the current state of the issue. The third button read “Pros and cons?” and clicking would open a balanced overview of no more than three arguments per side. These pro-con arguments were based on the arguments the political parties mentioned most frequently for and against each statement. We chose to always present an equal number of arguments pro and con, as users may otherwise be led by the number of arguments.

Besides using the suggestion buttons, users could request information by typing a question in the open chat window. An AI model was trained using open-source software Rasa to answer such user questions. In deciding which output to provide, the chatbot could recognize various intents: Semantic questions (“What is/are...?”), pragmatic questions (“How much...?”), reasons (“Why?”), party stances (“What does Y think?”), lead candidates (“Who’s the leader of Y?”), and all combinations of pros and cons (“What are the pros/cons/pros and cons”). Moreover, the chatbot could recognize VAA-related questions (“What does neither agree nor disagree mean?”), and it was able to respond to conversation-related expressions (reactions to “Hello”). Finally, the chatbot was also equipped with repair strategies in case it did not understand the question asked or in case the user said something rude.

While we used AI to recognize the user intent, no AI was used to generate the chatbot responses. Instead, the responses were written by humans. Information in the CA was based on reliable information resources (e.g., dictionaries and government websites) and was cross-checked by all members of the research team to

ascertain that the information was as correct and objective as possible (for a demonstration, see <https://surfdrive.surf.nl/files/index.php/s/UiuTsl3WDtdlCOq>).

3.3. Measures

3.3.1. Tool Evaluation Measures

Participants' evaluation of the tool was measured in a survey using a set of 7-point Likert-scales with questions measuring three aspects of the technology acceptance model (see Davis, 1989; Marangunić & Granić, 2015), namely ease of use (e.g., "I found the voting advice application easy to use"), usefulness (e.g., "This voting advice application was helpful"), and playfulness (e.g., "Using the voting advice application was fun"). These items were based on the items used in Kamoen and Liebrecht (2022; Study 2), who tailored the original survey questions of Ahn et al. (2007) to fit the context. The reliability of all scales was acceptable (ease of use: Cronbach's $\alpha = .79$; usefulness: $\alpha = .66$; playfulness: $\alpha = .84$).

3.3.2. Political Engagement Measures

We measured several different aspects of political engagement. First, participants' perceived political knowledge after tool usage was measured on a 7-point Likert-scale with three statements (e.g., "Using this voting advice application helps me better understand political parties' positions"). The same items were also used in the study of Kamoen and Liebrecht (2022, Study 1), who based these items on Ladner (2012). The reliability of this scale was acceptable only if the item that measured "the motivation to learn more about politics" was excluded ($\alpha = .74$).

Second, factual political knowledge was measured by letting participants judge seven statements about the political issues in the (CA)VAA, such as "In the Netherlands, currently no CO₂ is stored under the surface. True–false–don't know." For the analysis, an index of the number of correct answers was analyzed; "don't know" answers were scored as incorrect.

Third, internal political efficacy was measured using a single item: "I feel sufficiently informed to vote in the upcoming parliamentary elections." Finally, voting intention in terms of turnout was assessed with the item: "I am planning to vote in the elections of the Parliament on November 22, 2023." These questions were both based on Kamoen and Liebrecht (2022), who based these measures on Glynn et al. (2009).

3.3.3. Response Behavior

In both conditions, participants responded to the political attitude statements on a three-point scale, with an additional fourth "skip this question" option. Due to privacy regulations, we could not access participants' exact responses but were permitted to record whether each answer was directional (agree/disagree) or non-directional (neutral/skip). To compare the proportion of non-directional responses, we estimated a cross-classified multilevel logistic model with random effects for both items and participants (Quené & Van den Bergh, 2004, 2008). A contrast test (χ^2) was used to compare means in the fixed part.

3.4. Participants

Data were gathered between 31 October and 7 November 2023, about three weeks prior to the Dutch parliamentary elections on 22 November. We asked students from the participant pool of Tilburg University to participate in exchange for participant credits.

In total, 93 participants took part. Of them, 33 identified as male (32.4%) and 60 as female (66.9%). The average age was 20.89 years ($SD = 2.40$), ranging from 18 to 28. Almost all participants were students at the university level (97.8%), one participant preferred not to say (1.1%), and one was a student at the level of a higher vocational education (1.1%). All participants were allowed to vote in the upcoming elections. We compared the experimental groups for age and gender and found no differences (Gender: $\chi^2(1) = 0.31$, $p = .39$; Age: $t(91) = 0.77$, $p = .22$), which indicates that there is no reason to assume the groups differed *a priori*.

3.5. Procedure

Participants were seated in soundproof research cabins of the Dante laboratory at Tilburg University and were instructed on how to navigate between two tabs: one in which the tool (VAA or CAVAA) was open, and another tab that featured a survey. Participants started in the survey tab, where they read the instructions for how to fill out the (CA)VAA and, depending on the assigned condition, how to use the chatbot. After providing informed consent, they were instructed to go to the next tab in the browser. Here, they used the (CA)VAA, and after having received voting advice, participants were instructed to return to the previous tab to fill out the survey.

The survey measured tool evaluation and political engagement. Moreover, participants were also asked several other questions in the survey, including demographic variables and their general political interest. After completion, we informally asked the participants a couple of questions orally, mainly about their experience with the (CA)VAA. The purpose of this was to gather qualitative feedback on the workings of the tool. After having answered these interview questions, participants were thanked for their participation.

4. Results Study 1

Before discussing the results related to tool evaluation measures, political engagement measures, and the users' response behavior, we will first provide descriptive statistics related to the usage of the chatbot in the CAVAA.

4.1. Descriptive Statistics: Chatbot Usage

In the CAVAA condition, we monitored how participants ($N = 46$) used the chatbot. To quantify usage, we calculated the proportion of respondent-item combinations in which at least one suggestion button was clicked. Across all participants and items, the chatbot was used in 44.9% of the cases. The most frequently used button was "Pros and cons" (33.0%), followed by "What is the situation now?" (24.0%). The "What is X?" button, explaining terminology, was the least used (5.2%).

Participants submitted a total of 54 open-ended questions via the chatbot. In eight instances, multiple questions were asked by the same participant for the same VAA item, resulting in an overall open-chat usage rate of 3.3%. User questions were about various topics, and most of them related to party positions ($N = 16$). The chatbot provided correct responses to 48.1% of all queries.

4.2. Tool Evaluation and Political Engagement Measures

Table 1 shows the scores for the tool evaluation and political engagement measures for the VAA and the CAVAA.

Table 1. Means and standard deviations on a scale from 1 (*low score*) to 7 (*high score*) in Study 1.

Tool	Ease of use	Usefulness*	Playfulness	Perceived knowledge*	Factual knowledge	Internal political efficacy *	Turnout intention
CAVAA	6.49 (0.76)	5.93 (0.83)	5.51 (1.06)	4.68 (1.26)	3.11 (1.45)	4.78 (1.78)	6.63 (0.88)
VAA	6.18 (0.96)	5.48 (0.94)	5.52 (0.77)	4.12 (1.27)	3.02 (1.21)	3.72 (1.66)	6.38 (1.11)

Note: * = CAVAA > VAA with $p < .05$

As for tool evaluations, results show that the CAVAA is more useful than the VAA ($t(91) = 2.49$; $p = .014$, Cohen's $d = .51$). The size of this effect can be classified as medium-sized relative to the standard deviation. The differences for ease of use are not significant, although there is a tendency for an effect in the hypothesized direction with a small effect size ($t(91) = 1.68$; $p = .097$, Cohen's $d = .36$). No differences are observed for playfulness ($t(91) = 0.02$; $p = .99$).

Regarding the political engagement measures, results indicate that CAVAA users report higher levels of perceived political knowledge, and their internal political efficacy is also higher ($t(91) = 2.17$, $p = .03$, Cohen's $d = .44$; $t(91) = 2.97$, $p = .004$, Cohen's $d = .62$). The effects are small and medium-sized relative to the standard deviation. However, no significant differences are found for factual political knowledge or the extent to which participants were planning to vote ($t(91) = 0.32$, $p = .75$; $t(91) = 1.19$, $p = .24$).

4.3. Response Behavior

In the CAVAA condition, users provided non-directional answers in 13.6% of the respondent-item combinations, compared to 18.8% in the VAA condition. A contrast test indicates that this difference is statistically significant and large in terms of effect size: respondents gave fewer non-directional answers when using the CAVAA compared to the VAA ($\chi^2(1) = 4.33$, $p = .04$, Cohen's $d_{\text{items}} = 1.05$).

4.4. Summary

Results suggest that the CAVAA outperforms the VAA on several dimensions. In terms of tool evaluation, the CAVAA was rated higher on ease of use and usefulness, though not on playfulness. Regarding political engagement measures, CAVAA users reported higher perceived knowledge, and their internal political efficacy

was higher, but no significant differences were found in factual knowledge or turnout intention. Additionally, CAVAA users provided non-directional answers less frequently than VAA users.

5. Method Study 2

5.1. Design

Study 2 had a 2 (tool type: VAA or CAVAA) $\times$ 2 (user type: higher or lower levels of political sophistication) between-subjects design and was again conducted in the run-up to the 2023 Dutch parliamentary elections. This study aimed to replicate Study 1, but now with a more diverse sample of voters, and to explore whether the effects of VAAs versus CAVAAs are also dependent on the user's level of political sophistication. Respondents were randomly assigned to either the VAA or the CAVAA condition, and their level of political sophistication was measured as a quasi-experimental factor. The Research Ethics and Data Management Committee of approved the study (also under reference number REDC 2023.13a).

5.2. Materials and Measures

The same materials and measures as in Study 1 were used. Reliabilities were acceptable for all DVs that were measured with more than one item (ease of use: $\alpha = .82$; usefulness: $\alpha = .79$; playfulness: $\alpha = .88$; perceived political knowledge, 2 items: $\alpha = .82$).

To distinguish between participants with relatively higher and lower levels of political sophistication, we combined educational level and political interest into an aggregate index (Kamoen & Liebrecht, 2022). Political interest was measured using three items ($\alpha = .89$) based on Lachat (2008). The mean interest score was rescaled to a 4-point scale to match the educational scale, ensuring equal weight. We then averaged the two measures and performed a median split (cut-off = 2.50) to classify participants into higher and lower political sophistication groups.

5.3. Participants

A total of 149 participants took part in Study 2. For one of these participants, we were unable to trace whether the participant filled out the tool, and these data were therefore not considered for the analyses. Four other participants did not fill out any demographic questions in the survey. Because two such demographic questions were used to construct an indicator for political sophistication, it was not possible to determine political sophistication for these respondents, and their data were consequently excluded from the analyses. Of the remaining participants ($N = 144$), 74 identified as male (51.4%), 64 as female (44.4%), and six participants did not identify as either male or female (4.2%). The average age was 33.12 years ($SD = 15.85$).

We checked the gender and age distribution across the tool types and the two groups based on the level of political sophistication. No differences were observed in mean age (in all cases: $F(1, 139) < 1.28$; $p > .26$), and the male/female ratio was comparable in these groups (in both cases: $\chi^2(1) < .23$; $p > .63$).

5.4. Procedure

This study was conducted outside the lab between 9 November and 22 November 2023. Participants were approached in person by one of the researchers in several locations, including hairdressers, tattoo shops, and municipal buildings. In all these locations, we had consent to approach people, and if participants agreed to take part, they were seated in a quiet location within the building. From that point onwards, the procedure was the same as the procedure followed in Study 1.

6. Results Study 2

6.1. Tool Evaluations and Political Engagement Measures

Table 2 shows the means and standard deviations per condition. For both usefulness and playfulness, a medium-sized main effect of tool type is observed (usefulness: $F(1, 140) = 13.79, p < .001$, Cohen's $d = .63$; playfulness: $F(1, 140) = 8.96, p = .003$, Cohen's $d = .53$). Participants perceive the CAVAA as more useful and playful than the regular VAA. These effects are consistent across user groups, as both the main effect of sophistication and the interaction between tool type and sophistication were not significant (usefulness: $F(1, 140) = 0.21, p = .62$ and $F(1, 140) = 0.03, p = .86$; playfulness: $F(1, 140) = 0.33, p = .57$ and $F(1, 140) = 0.05, p = .83$). In line with findings in Study 1, the effect for ease of use is only marginally significant ($F(1, 140) = 3.12, p = .08$), which is consistent again across user groups since no significant main or interaction effect was found ($F(1, 140) = 0.20, p = .90$; $F(1, 140) = 0.77, p = .38$).

Table 2. Means and standard deviations on a scale from 1 (low score) to 7 (high score) in Study 2.

Political sophistication	Tool	Ease of use	Usefulness*	Playfulness*	Perceived knowledge*	Factual knowledge ⁺	Internal political efficacy ⁺	Turnout intention ⁺
Low	CAVAA	6.38 (0.76)	5.67 (0.94)	5.49 (1.16)	4.79 (1.35)	2.69 (1.75)	4.75 (1.78)	6.06 (1.64)
	VAA	5.90 (1.22)	5.02 (1.23)	4.97 (1.28)	4.36 (1.32)	2.37 (1.80)	4.49 (1.76)	5.97 (1.52)
High	CAVAA	6.25 (1.41)	5.79 (1.20)	5.64 (0.96)	4.94 (1.60)	4.08 (1.36)	5.42 (1.65)	6.87 (0.53)
	VAA	6.09 (0.75)	5.08 (0.99)	5.04 (1.09)	4.06 (1.33)	3.94 (1.41)	5.14 (1.77)	6.69 (0.76)

Notes: ⁺ = Main effect of political sophistication: higher scores for higher politically sophisticated users, * = main effect of tool type: higher scores in the CAVAA condition compared to the VAA condition.

As for the political engagement measures, results show a small-sized main effect of tool type for perceived knowledge ($F(1, 140) = 7.92, p = .006$, Cohen's $d = .47$), with scores being higher for CAVAA users. This effect is consistent across participants' levels of political sophistication, as neither the main effect of sophistication ($F(1, 140) = 0.00, p = .995$) nor the interaction reached statistical significance ($F(1, 140) = 0.71, p = .40$).

Factual knowledge scores do not differ between the CAVAA and VAA ($F(1, 138) = 0.72, p = .40$). However, results show a large significant effect of political sophistication: participants with lower levels of political sophistication score substantially lower on factual knowledge compared to their more sophisticated peers

($F(1, 138) = 30.81, p < .001$, Cohen's $d = 1.00$). This difference does not appear to be driven by tool usage, as the interaction between tool type and sophistication was not significant ($F(1, 138) = 0.72, p = .40$).

Finally, results for internal political efficacy and turnout intention show a consistent pattern: no significant differences appear between tool types ($F(1, 140) = 0.88, p = .35$; $F(1, 140) = 0.72, p = .40$). In contrast, political sophistication has a significant effect on these scores, as more politically sophisticated participants report higher internal efficacy scores and also report being more likely to vote ($F(1, 140) = 5.25, p = .02$, Cohen's $d = .38$; $F(1, 140) = 14.48, p < .001$, Cohen's $d = .69$). Again, these effects are not moderated by tool type, as no significant interactions are found (self-efficacy: $F(1, 140) = 0.001, p = .98$; turnout intention: $F(1, 140) = 0.11, p = .74$).

6.2. Response Behavior

Regarding the proportion of non-directional responses (see Table 3), results show tendencies for main effects: both lower politically sophisticated users ($\chi^2(1) = 3.67, p = .056$) and VAA users are more likely to provide non-directional responses ($\chi^2(1) = 2.88, p = .09$). The interaction between tool type and respondent type is not statistically significant ($\chi^2(1) = 0.11, p = .75$). The differences between conditions and levels of political sophistication are large relative to the between-item standard deviation (Cohen's $d > 1.23$) and also large in terms of percentages (roughly a 4% difference between tool types and user groups).

Table 3. Percentage non-directional answers for CAVAA and VAA users with higher and lower political sophistication in Study 2.

Political sophistication	Tool	% non-directional answers (Logit)	S ² items	S ² participants
Low	CAVAA	15.2% (−1.72)	0.199	0.67
	VAA	20.0% (−1.39)		0.40
High	CAVAA	12.0% (−1.98)	0.97	0.97
	VAA	14.7% (−1.76)		0.85

6.3. Chatbot Usage

Table 4 displays chatbot usage by user type. Results indicate that less politically sophisticated participants more frequently requested semantic information ($\chi^2(1) = 9.44, p = .003$, Cohen's $d_{\text{items}} = 1.09$), while higher politically sophisticated users more often consulted pro and con arguments ($\chi^2(1) = 12.95, p < .001$, Cohen's $d_{\text{items}} = 1.04$). No significant difference was found for the use of pragmatic information ($\chi^2(1) = 1.13, p = .32$). This suggests that the type of information users seek varies with their level of political sophistication, even though the overall proportion of users requesting information via buttons at least once did not differ between these groups ($\chi^2(1) = 0.53, p = .47$).

In addition to button usage, participants submitted a total of 67 questions via open chat. After accounting for multiple open-ended questions related to the same respondent and item, this corresponds to 2.3% of respondent-item combinations. Including these open-ended questions in the overall measure of information requests, information was accessed in approximately 24% of cases, with no significant differences between more and less politically sophisticated users.

Table 4. Proportion of respondent-item combinations in which a specific button is clicked at least once in Study 2 (Logit between brackets).

Political sophistication	Semantic information button*	Pragmatic information button	Pro and con argument button*	Clicking at least 1 button	Requesting information (incl. chat)
Low	4.3% (–3.09)	10.5% (–2.14)	13.9% (–1.82)	24.8% (–1.19)	24.8% (–1.19)
High	1.8% (–3.88)	9.1% (–2.29)	19.8% (–1.40)	23.4% (–1.11)	23.4% (–1.12)

Note: * = $p < .05$

6.4. Summary

The overall positive effects of the CAVAA compared to the VAA, as found in Study 1, are also observed in a more diverse sample. Moreover, Study 2 shows that tool type does not interact with user type across tool evaluation measures, political engagement measures, or response behavior to VAA statements. Regarding information-seeking behavior within the CAVAA, differences between lower and higher politically sophisticated users emerged: less sophisticated users more frequently request explanations of terms, while more sophisticated users more often review pro and con arguments. Given these different behavioral patterns, a relevant question that will be addressed in Study 3 is how both users with high and low political sophistication evaluate and interact with a CAVAA compared to an enhanced VAA (VAA+) that includes supplementary information in the web environment of the tool.

7. Method Study 3

7.1. Design

Study 3 employed a 2 (tool type: VAA+ or CAVAA) $\times$ 2 (user type: higher or lower levels of political sophistication) between-subjects design, comparable to the design in Study 2. Participants were randomly assigned to one of the two tool types, while political sophistication was measured quasi-experimentally. The study was conducted in the lead-up to the European parliament election in the Netherlands on 6 June 2024. Ethical approval was obtained from the Research Ethics and Data Management Committee of the Tilburg School of Humanities and Digital Sciences at Tilburg University (reference number: REDC 2023.13b).

7.2. Materials

As in Studies 1 and 2, we developed two versions of the tool based on a copy of StemWijzer (Figure 2). The tools featured 30 political attitude statements relevant to the 2024 European parliament elections. In the VAA+ version, users could access four types of additional information within the web interface. Semantic explanations of difficult terms were available by clicking on highlighted words in the statements. Information about the status quo, pro and con arguments, and party stances were accessible through a button.

In the CAVAA condition, a chatbot was integrated into an otherwise empty version of the VAA, without highlighted terms or clickable buttons. To maintain functional equivalence with the VAA+, the chatbot included four preset buttons corresponding to the same information types. Additionally, it featured an open

text field, allowing users to type their own questions and access a wider range of information than the static VAA+ interface. The chatbot was built using the same approach as in previous studies.

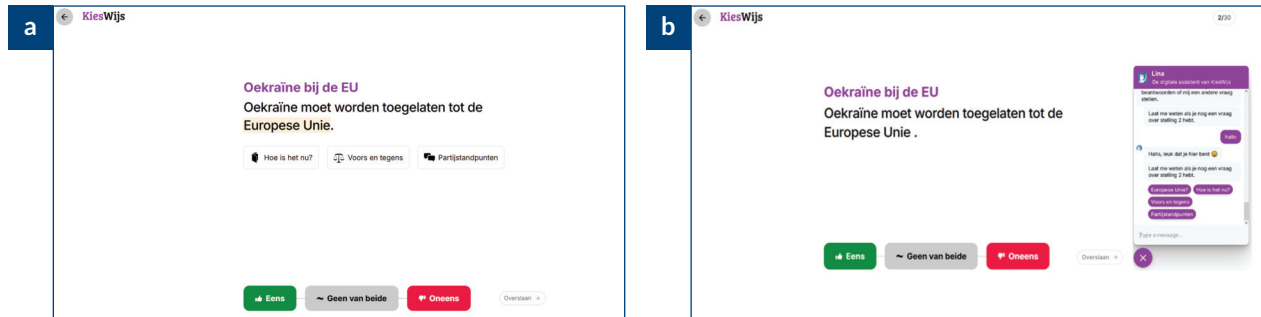


Figure 2. Screenshot of the VAA+ (a) and CAVAA (b). Notes: The header reads “Ukraine’s accession to the EU” and the statement is “Ukraine should join the EU.” In the VAA+, the highlighted term (“European Union”) is clickable, and three additional clickable buttons are displayed below the statement; in the CAVAA, four buttons are displayed, each leading to the same information as in the VAA+ (i.e., via the three buttons and the highlighted term).

7.3. Measures

Political sophistication was measured as in Study 2 by combining political interest ($\alpha = .82$) and education. Tool evaluation and political engagement measures were identical to previous studies, adapted to the European election context. Just like in Study 1 and 2, the scales showed acceptable consistency (ease of use: $\alpha = .93$; usefulness: $\alpha = .76$; playfulness: $\alpha = .80$; perceived knowledge: 2 items; $\alpha = .87$). We again measured users’ factual political knowledge, logged directional versus non-directional responses, and tracked button clicks and open chat questions in the CAVAA condition. Differences between conditions were analyzed using multilevel models, following the approach from Studies 1 and 2.

7.4. Participants and Procedure

Study 3 was conducted outside the lab between 22 May and 6 June 2024, following a comparable procedure to that used in Study 2. A total of 159 participants took part: 92 participants identified as male (57.9%), 66 as female (41.5%), and one as other (0.6%). The average age was 45.06 years ($SD = 19.63$). We compared gender and age distribution across conditions. No differences were observed in mean age (in all cases: $F(1, 139) < 1.28$; $p > .26$) and gender ($\chi^2(1) < .63$; $p > .73$).

8. Results Study 3

8.1. Tool Evaluation and Political Engagement Measures

The scores for ease of use, usefulness, and playfulness do not differ between the two tool types, nor between participants with higher or lower levels of political sophistication (in all cases: $F(1, 155) < 2.62$, $p > .11$; see Table 5).

For the political engagement measures, only the main effects of political sophistication are observed. Consistent with findings from Study 2, more politically sophisticated users' internal self-efficacy is higher, they report a higher turnout intention and score higher on factual knowledge questions (in all cases: $F(1, 155) > 5.28, p < .02$). No other main effects or interactions between tool type and political sophistication are observed (in all cases: $F(1, 155) < 1.36, p > .25$).

Table 5. Means and standard deviations on a scale from 1 (*low score*) to 7 (*high score*) in Study 3.

	Ease of use	Usefulness	Playfulness	Perceived knowledge	Factual knowledge	Internal political efficacy	Turnout intention
VAA+, low sophistication	6.27 (0.55)	5.63 (0.83)	5.41 (1.01)	4.96 (1.10)	2.44 (1.36)	5.17 (1.48)	5.72 (1.70)
VAA+, high sophistication	6.45 (1.03)	5.73 (1.02)	5.68 (0.91)	5.10 (1.21)	3.58 (1.31)	5.60 (1.10)	6.56 (1.04)
CAVAA, low sophistication	6.36 (1.06)	5.89 (0.97)	5.60 (1.10)	4.72 (1.56)	2.74 (1.85)	5.38 (1.35)	5.56 (1.79)
CAVAA, high sophistication	5.99 (1.43)	5.68 (1.06)	5.66 (1.07)	4.85 (1.40)	3.31 (1.67)	5.87 (1.11)	6.44 (1.19)

8.2. Response Behavior

The proportion of non-directional answers (Table 6) did not differ between participants with relatively high versus low levels of political sophistication, between tool types, or in their interaction (in all cases: $\chi^2(1) < 0.82, p > .37$).

Table 6. Percentage non-directional answers in Study 3 (Logit used in the analyses between brackets).

	Non-directional answers	Semantic information button	Pragmatic information button	Pro and con arguments button	Party stances button
VAA+, low sophistication	11.1% (−2.08)	1.3% (−4.32)	13.3% (−1.88)	21.6% (−1.29)	6.6% (−2.64)
VAA+, high sophistication	10% (−2.20)	2.0% (−3.89)	12.7% (−1.92)	17.6% (−1.54)	15.4% (−1.70)
CAVAA, low sophistication	11.9% (−2.00)	5.6% (−2.82)	7.6% (−2.49)	11.9% (−2.00)	3.8% (−3.22)
CAVAA, high sophistication	10.1% (−2.19)	2.5% (−3.67)	4.3% (−3.11)	6.3% (−2.70)	2.9% (−3.51)

8.3. Information Requests Made

There was a tendency for semantic information to be requested more often in the CAVAA than in the VAA+ ($\chi^2(1) = 3.15, p = .08$), but neither political sophistication ($\chi^2(1) = 0.19, p = .66$) nor its interaction with tool type ($\chi^2(1) = 1.77, p = .32$) was significant.

For other information types, a clear pattern emerged: pragmatic information, pro and con arguments, and party stances were all requested significantly more often in the VAA+ than in the CAVAA (in all cases $\chi^2 > 7.49$,

$p < .01$), with no significant effects for political sophistication or its interaction (all χ^2 s < 2.74 , p 's $> .16$). The size of the observed differences between tool types is consistently large relative to the between-item variance, and also substantial in practical terms, as some information types are requested twice or even three times as often in the VAA+ condition (Table 6).

Among the 78 CAVAA users, 40 open-ended questions were submitted (Response accuracy: 27.5%), corresponding to 27 unique respondent-item combinations (1.15%). Including these in the total measure of information-seeking behavior confirms the same overall pattern: participants requested information more frequently in the VAA+ than in the CAVAA.

8.4. Summary

No differences can be observed between the VAA+ and CAVAA in tool evaluation, political engagement measures, or the proportion of non-directional answers—regardless of political sophistication. Differences do emerge in the information requests made. There is tendency for semantic information to be requested more often in the CAVAA, while status quo information, pro and con arguments, and party stances are more frequently accessed in the VAA+.

9. General Discussion

In three experimental studies, we examined the effects of CAVAA compared to traditional VAAs without additional information (Studies 1 and 2), and to an enhanced VAA with clickable information buttons (VAA+; Study 3). The studies differed in sampling strategy and context: Study 1 was conducted in a controlled lab setting with a sample of highly educated participants, whereas Studies 2 and 3 targeted more diverse populations, allowing for a comparison between groups in their level of political sophistication. In terms of timing, Studies 1 and 2 took place in the run-up to the 2023 Dutch parliamentary elections (first-order elections), while Study 3 was conducted during the 2024 European elections (second-order elections). In the next sub-sections, we reflect on the most important findings across these studies and relate the findings to theory and practice.

9.1. VAAs versus CAVAA

Studies 1 and 2 show that CAVAA users actively engaged with the chatbot, using buttons in 25% (Study 2) to 43% (Study 1) of respondent-item combinations, while typing questions in 2–3% of the cases. This supports Krosnick's (1991) satisficing-optimizing model: the less effort required to access information, the more likely voters are to engage. In other words, easily accessible information encourages voters to read it, which aligns well with the core purpose of VAAs—to inform voters.

CAVAAs were also evaluated more favorably than VAAs in terms of usefulness and perceived knowledge (Studies 1 and 2), internal political efficacy (Study 1), and perceived playfulness (Study 2). Despite this variation in the exact effects observed, the overall pattern is clear and in line with H1 and H2, CAVAA outperform traditional VAAs without information in terms of tool evaluation and political engagement measures. This conclusion aligns with Study 1 in Kamoen and Liebrecht (2022). Notably, our studies used more realistic and construct-valid materials.

As in Kamoen and Liebrecht (2022), current studies did not indicate an effect of tool type on the extent to which users plan to cast a vote. Kamoen and Liebrecht (2022) attributed this to the lack of ecological validity in their study, which is not an explanation here. In current studies, the scores for voting intention were consistently high (6.5 on a 7-point scale), which might point to a ceiling effect and perhaps to providing socially desirable answers. A suggestion for future research is therefore to construct a survey item playing into this potential bias.

Current studies also included the response behavior to political attitude statements as an unobtrusive indicator of comprehension difficulties. CAVAA users provided fewer non-directional responses compared to users of traditional VAAs, although the effect in Study 2 was only marginally significant. Hence, in line with H3, providing additional information in a chatbot is likely to mitigate comprehension challenges and satisficing behavior when filling out a VAA.

9.2. VAA+ Versus CAVAA

A question is whether the benefits of the CAVAA stem from the chatbot interface itself or from the availability of additional information. Results of Study 3 showed no significant differences between CAVAA and VAA+ in terms of tool evaluations, political engagement measures, or response behavior. This suggests that the presence of information, regardless of how it is accessed, is the main driver of the effects observed in Studies 1 and 2.

Interestingly, Study 3 did reveal differences in the type and frequency of information accessed through the CAVAA versus the VAA+, as there was a tendency for semantic information to be requested more frequently in the CAVAA, whereas status quo information, pro and con arguments, and party stances were more frequently accessed in the VAA+. These findings raise several points for discussion. First, why would one type of information be requested more frequently in the CAVAA, whereas all other types are more often consulted in the VAA+? A likely explanation lies in the way information was presented. In the VAA+, semantic information was available via highlighted terms within the statements, while other types of information were accessible via buttons. Post-study interviews indicated that some users interpreted the highlighted terms as indicators of importance rather than as interactive elements, potentially explaining the lower access rate. It would be worthwhile to test this explanation in a future study.

Another point concerns how to interpret the finding that information was accessed more frequently in the VAA+ than in the CAVAA. One explanation is that the VAA+ aligns better with the low-effort information processing strategies (Krosnick, 1991) that many users apply. Alternatively, VAA+s might facilitate information access more effectively because users are more familiar with this interface. If so, we may expect differences in usage patterns between CAVAAs and VAA+s to diminish over time as users are expected to become more accustomed to conversational interfaces. A third interpretation is that VAA+ users may have clicked multiple buttons in search of specific information, while CAVAA users were able to formulate more targeted queries through open chat. In addition, in the CAVAA it was possible to re-read texts that were requested, as it was always possible to scroll through the chat, whereas an additional button-click was required to re-read button-information from the VAA+. If this explanation is valid, the higher number of clicks in the VAA+ condition might underscore the potential value of developing CAVAAs. Future research should explore these mechanisms in greater depth to better understand the differing engagement patterns across formats.

9.3. Political Sophistication

In Studies 2 and 3, we distinguished users based on their level of political sophistication. A key finding is that tool type effects were not moderated by political sophistication, suggesting that the benefits of contextual information are consistent across user groups. The studies did reveal several main effects of sophistication: more politically sophisticated users scored higher on factual knowledge and reported a higher likelihood of voting. These outcomes align with expectations of what one might expect for more and less politically sophisticated users, and therefore merely give confidence in the construct validity of the measurement.

In Study 2, we also observed differences in information-seeking behavior: less politically sophisticated users more often accessed semantic information, whereas more sophisticated users more frequently consulted pro and con arguments. Interestingly, these patterns did not emerge in Study 3. We attribute this discrepancy to the nature of the election, and more specifically, to how this may have affected the sample we were able to recruit. While Studies 1 and 2 were conducted during first-order national elections, Study 3 focused on the second-order European parliament elections, which typically receive less public attention, media coverage, and voters also tend to know less about such elections (Reif & Schmitt, 1980). We suspect that differences in public interest between first- and second-order elections affected not only who participated in the study, but also the nature of the participant groups. Although recruitment took place in similar locations, recruiting participants during the European elections was noticeably more difficult. The samples from Studies 2 and 3, therefore, differed, not just in composition; for instance, participants in Study 3 were generally older, but may have also differed in the underlying characteristics of the higher and lower-sophisticated groups.

10. Conclusion

In conclusion, our findings suggest that embedding contextual assistance in a VAA, whether through a chatbot or clickable buttons in the web environment, improves tool evaluation and political engagement measures, and also helps users provide directional answers to the statements. These benefits occur across levels of political sophistication, making both VAA+s and CAVAAs promising tools to encourage users to inform themselves about the political issues at stake. At the same time, more research is needed to understand why voters request more information in a VAA+ than in a CAVAA. Based on current findings, it is not possible to determine whether this is an argument in favor of CAVAA or VAA+s. We therefore encourage further research on the integration of (controlled) chatbots in the context of VAAs, but in a way that balances both the information needed and the ethical considerations of voters, governments, and supra-national agreements such as the AI Act.

Acknowledgments

We thank ProDemos for granting us the right to use a copy of their tool. We would like to thank Elke van Veggel, Lotte Tiebosch, and Doris Verbunt for their assistance with collecting the data. We are also very grateful to Diana Schmalzried for proofreading our manuscript.

Funding

We received funding from our department for conducting Studies 1 and 2. We want to thank our head of department for this. Additionally, for Study 3, we received funding from ProDemos for developing the CA and for hiring student assistants for data collection. Publication of this article in open access was made possible through the institutional membership agreement between Tilburg University and Cogitatio Press.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The surveys and data files of all studies can be found via <https://doi.org/10.34894/R2YWEL>

LLMs Disclosure

The authors used GPT-4o solely to rephrase sentences for conciseness and grammatical correctness.

References

- Ahn, T., Ryu, S., & Han, I. (2007). The impact of web quality and playfulness on user acceptance of online retailing. *Information & Management*, 44(3), 263–275. <https://doi.org/10.1016/j.im.2006.12.008>
- Baka, A. S., Figgou, L., & Triga, V. (2012). ‘Neither agree nor disagree.’ A critical analysis of the middle response option in voting advice applications. *International Journal of Electronic Governance*, 5(3/4), 244–263.
- Bruins Slot, H. (2022, October 10). *Uitkomsten onderzoek naar opkomst bij gemeentereedsverkiezingen 2022*. Rijksoverheid. https://www.tweedekamer.nl/kamerstukken/brieven_regering/detail?id=2022D40761&did=2022D40761
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13, 319–340. <https://doi.org/10.2307/249008>
- De Graaf, J. (2010). The irresistible rise of Stembijzetter. In L. Cedroni & D. Garzia (Eds.), *Voting advice applications in Europe. The state of the art* (pp. 35–60). ScriptaWeb.
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Dieing, T. I. (2025). AI-driven dialogue: Leveraging generative AI in conversational agent voting advice applications (CAVAAs). In A. Følstad, S. Papadopoulos, T. Araujo, E. L.-C. Law, E. Luger, S. Hobert, & P. B. Brandtzaeg (Eds.), *Chatbots and human-centered AI* (pp. 161–180). Springer. https://doi.org/10.1007/978-3-031-88045-2_11
- Elling, S., Lentz, L., & De Jong, M. (2012). Users’ abilities to review web site pages. *Journal of Business Technical Communication*, 26(2), 171–201. <https://doi.org/10.1177/1050651911429920>
- Fivaz, J., & Nadig, G. (2010). The impact of voting advice applications (VAAs) on voter turnout and their potential use for civic education. *Policy & Internet*, 2(4), 167–200. <https://doi.org/10.2202/1944-2866.1025>
- Garzia, D., & Marschall, S. (2012). Voting advice applications under review: The state of research. *International Journal of Electronic Governance*, 5, 203–222.
- Glynn, C. J., Huge, M. E., & Lunney, C. A. (2009). The influence of perceived social norms on college students’ intention to vote. *Political Communication*, 26(1), 48–64. <https://doi.org/10.1080/10584600802622860>
- Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 213–236.
- Hankel, S., Liebrecht, C., & Kamoen, N. (2024). ‘Hi chatbot, let’s talk about politics!’ Examining the impact of verbal anthropomorphism in Conversational Agent Voting Advice Applications (CAVAAs) on higher and lower politically sophisticated users. *Interacting With Computers*, 37(4), 221–233. <https://doi.org/10.1093/iwc/iwae031>
- Heinsohn, T., Fatke, M., Israel, J., Marschall, S., & Schultze, M. (2019). Effects of voting advice applications during election campaigns. Evidence from a panel study at the 2014 European elections. *Journal of Information Technology & Politics*, 16(3), 250–264. <https://doi.org/10.1080/19331681.2019.1644265>

- Kamoen, N., & Holleman, B. (2017). I don't get it. Response difficulties in answering political attitude statements in voting advice applications. *Survey Research Methods*, 11(2), 125–140. <https://doi.org/10.18148/srm/2017.v11i2.6728>
- 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.
- Kamoen, N., & Liebrecht, C. (2022). I need a CAVAA: How conversational agent voting advice applications (CAVAAs) affect users' political knowledge and tool experience. *Frontiers in Artificial Intelligence*, 5, Article 835505. <https://doi.org/10.3389/frai.2022.835505>
- Kamoen, N., Liebrecht, C., Van Lieshout, R., & Van Veggel, E. (2024, August 12–14). "Hi voter, I can assist you in filling out this voting advice application!" How CAVAAs may assist voters in informing themselves about elections [Paper presentation]. ECPR General Conference, Dublin, Ireland.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. <https://doi.org/10.1002/acp.2350050305>
- Lachat, R. (2008). The impact of party polarization on ideological voting. *Electoral Studies*, 27(4), 687–698. <https://doi.org/10.1016/j.electstud.2008.06.002>
- Ladner, A. (2012, April 3–5). *Voting advice applications—Impact on voting decisions in the 2011 Swiss national elections* [Paper presentation]. The 62nd Political Studies Association Annual International Conference, Belfast, Northern Ireland. https://www.researchgate.net/publication/261495685_Voting_advice_applications--impact_on_voting_decisions_in_the_2011_Swiss_national_elections
- Luskin, R. C. (1990). Explaining political sophistication. *Political Behavior*, 12(4), 331–361. <https://doi.org/10.1007/BF00992793>
- Marangunić, N., & Granić, A. (2015). Technology acceptance model: A literature review from 1986 to 2013. *Universal Access in the Information Society*, 14, 81–95. <https://doi.org/10.1007/s10209-014-0348-1>
- Mason, L. E., Krutka, D. G., & Stoddard, J. (2018). Media literacy, democracy, and the challenge of fake news. *Journal of Media Literacy Education*, 20(2), 1–10.
- Munzert, S., & Ramirez-Ruiz, S. (2021). Meta-analysis of the effects of voting advice applications. *Political Communication*, 38(6), 691–706. <https://doi.org/10.1080/10584609.2020.1843572>
- Quené, H., & Van den Bergh, H. (2004). On multilevel modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43, 103–121.
- Quené, H., & Van den Bergh, H. (2008). Examples of mixed effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59, 413–442.
- Reif, K., & Schmitt, H. (1980). Nine second-order national elections: A conceptual framework for the analysis of European election results. *European Journal of Political Research*, 8(1), 3–44.
- Roberts, C., Gilbert, E., Allum, N., & Eisner, L. (2019). Research synthesis. Satisficing in surveys: A systematic review of the literature. *Public Opinion Quarterly*, 83(3), 598–626. <https://doi.org/10.1093/poq/nfz035>
- Sipma, T., Lubbers, M., Van der Meer, T. W., Spierings, D., & Jacobs, K. (Eds.). (2021). *Versplinterde vertegenwoordiging: Nationaal Kiezersonderzoek 2021*. Stichting KiezersOnderzoek Nederland. https://dspace.library.uu.nl/bitstream/handle/1874/421298/NKO_2021_Versplinterde_vertegenwoordiging.pdf?sequence=1
- Van de Pol, J., Holleman, B. C., Kamoen, N., Krouwel, A. P. M., & De Vreese, C. H. (2014). Beyond young, higher educated males: A typology of VAA users. *Journal of Information, Technology and Politics*, 11, 397–411. <https://doi.org/10.1080/19331681.2014.958794>
- Van Outersterp, R., Kamoen, N., & Holleman, B. (2016). Inwisselbaar of niet? De middelste antwoordoptie en de non-responsoptie in stembulpen. *Tijdschrift voor Taalbeheersing*, 38(1), 27–55.

- Van Veggel, E., Liebrecht, C., & Kamoen, N. (2025). How life-like digital humans in voting advice applications can stimulate young voters to inform themselves about politics: The role of familiarity and expertise. *International Journal of Human-Computer Interaction*, 41(20), 12816–12830. <https://doi.org/10.1080/10447318.2025.2465859>
- Van Zanten, N., & Boumans, R. (2023). Voting assistant chatbot for increasing voter turnout at local elections: An exploratory study. In A. Folstad, T. Araujo & S. Papadopoulos (Eds.), *International Workshop on Chatbot Research and Design* (pp. 3–22). Springer. https://doi.org/10.1007/978-3-031-54975-5_1
- Wall, M., Krouwel, A. P. M., & Vitiello, T. (2014). Do voters follow the recommendations of voter advice application websites? A study of the effect of kieskompas.nl on its users' vote choices in the 2010 Dutch Legislative Elections. *Party Politics*, 20(3), 416–428. <https://doi.org/10.1177/135406881143605>

About the Authors



Naomi Kamoen is an assistant professor at Tilburg University. Her current research focuses on the uses and effects of VAAs. She collaborates with societal partners in real-life election settings, using a wide range of research methods such as experiments, content analyses, and big data analyses.



Christine Liebrecht is an assistant professor in the Department of Communication and Cognition at Tilburg University. Her research examines the transformation of human communication through technology across different user types and contexts. Her approach incorporates both quantitative and qualitative research methodologies, frequently linking insights to practical applications.



Rieke van Lieshout is a research master's student in linguistics at Radboud University Nijmegen. Her current research focuses on simple language and on how large language models can improve (linguistic) accessibility. Combining computational and theoretical expertise, she builds bridges across disciplines and in society.

ARTICLE

Open Access Journal 

Party Competition and Voter Attitudes in German Border Regions: Evidence From Local VAAs

Daniela Braun ¹ , Elisa Deiss-Helbig ² , Theresa Gessler ³ , Jochen Müller ⁴ ,
Julia Wagner ¹ , and Georg Wenzelburger ¹ 

¹ Department of European Social Research, Saarland University, Germany

² Department of Politics and Public Administration, University of Konstanz, Germany

³ Department of Social Sciences, University of Hamburg, Germany

⁴ Department of Political Science and Communication Studies, University of Greifswald, Germany

Correspondence: Georg Wenzelburger (georg.wenzelburger@uni-saarland.de)

Submitted: 28 August 2025 **Accepted:** 24 November 2025 **Published:** 28 January 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

When studying questions of European democracy, research is still very much focused on national political actors. This can be partly explained by the fact that data on political parties' positions toward Europe and citizens' attitudes have been mainly collected at the national level, and because sample sizes at the regional or local levels are too small. However, political parties compete about issues related to Europe on different levels—and especially so in European border regions where the local level coincides even more strongly with the EU level. For the particular case of such inner-EU border regions, home to around one third of EU citizens, little is known about patterns of party competition and citizens' specific demands and preferences. This study addresses this gap, using innovative data from the voting advice application VOTO, specifically designed for local elections held in Germany in 2024. Focusing on four German border regions—two on the Eastern and two on the Western border—it provides new insights into how European democracy works at the local level by examining political parties' positions and citizens' political preferences. In particular, the study investigates the extent to which border-specific features and party cues shape voters' perspectives on cross-border cooperation and European integration. Our findings speak not only to insights from qualitative case studies, which tell us that citizens living in such border regions have specific preferences in terms of saliency and positions, but also to quantitative research studying Euroscepticism in European border regions.

Keywords

border regions; Europe; Germany; party competition; political attitudes; VAAs

1. Introduction

Research on the transformation of politics in European democracies has mainly been focused on the level of the EU and its connections with the nation-states (De Vries & Hobolt, 2020; Hooghe & Marks, 2018; Hutter et al., 2016; Scharpf, 2019). Over the last decades, this vibrant field has produced important insights about the rise of right-wing populism in Europe and second-dimension politics (Bornschieer, 2010), its relationship to Euroscepticism (Treib, 2021; Vasilopoulou & Zur, 2024), and—although less extensively—the repercussions for policy-making at the national and European levels (Debus & Schulte, 2024; Thomas & Jones, 2025). However, as recent scholarship has demonstrated, looking at subnational levels of government allows us to uncover some of the dynamics that underpin current political transformations in Europe. Schraff and Pontusson (2024) show, for instance, how the structure of territorial inequalities influences voting for right-wing populist parties, and similar results have been delivered by Vasilopoulou and Talving (2024) regarding trust in the EU. Still, this emergent literature is mostly based on data on voting behaviour and cannot fully account for the supply side of party competition, namely, how political parties position themselves on European issues on a subnational level (for an exception, see Gross, 2022; Gross & Debus, 2018). This is mainly due to the fact that fine-grained data on party positions, electoral behaviour, and elections on subnational levels is scarce (see Röth et al., 2025), and data collection at the local level is demanding (Gross & Jankowski, 2020). Therefore, most research questions that aim to explore how parties compete for votes on Europe-related issues in specific regional contexts cannot easily be answered.

In this article, we focus on European border regions as the most relevant example of such a specific regional context. To gain insights for such a rarely studied context, we draw on data from voting advice applications (VAAs). Although VAAs are mainly used as tools to help voters find their best-matching candidates or parties (e.g., Dinas et al., 2014), this is not the only way in which data generated through VAAs can be exploited. As Garzia and Marschall (2014) rightly point out, VAAs can deliver interesting data to identify political dimensions (Wheatley, 2016), to study in which way political parties translate their positions into policy-making once they are in power (Fivaz et al., 2014), or to investigate the making of a transnational European voting space (Bright et al., 2016). We show that VAA data can also be efficiently used to study research questions about local—or regional-level problems in a much more fine-grained way. The VOTO VAA (Stecker et al., 2025) on which we base our study was fielded in 2024 at the local government level in 26 German municipalities, including the four border regions analysed in this article, and was followed by a short survey to collect data on attitudes of participants. The data gathered through VOTO allows us to study the following three research questions: First, we examine how political parties position themselves at the local level of border regions toward issues linked to cross-border cooperation and European integration. Second, we investigate whether cross-border issues and EU issues are correlated with the level of VAA participants. And, finally, we explore to what extent participants' attitudes, such as party preferences or ideology, are related to their respective positions on cross-border and EU issues. Zooming in on border regions is particularly promising, as European topics—such as border controls or the common market—are part of people's everyday lives. Moreover, the debate on whether citizens living in border regions are more or less Eurosceptic (Nasr & Rieger, 2024; Rehm et al., 2025) and have more or less transnational attitudes (Kuhn, 2011) shows that focussing on these interconnected areas in which around 30 percent of EU citizens live can deliver relevant insights on how parties and voters compete on EU issues at the local level.

The remainder of the article is structured as follows. We first provide insights into our theoretical argument that border regions in Europe are exceptional regions that may affect individuals' political preferences. In view of the fact that appropriate data is not available to study these political preferences in these contexts, we describe in a subsequent step how we applied data from the VOTO VAA, specifically designed for local elections held in Germany in 2024, for the study of our research questions. Next, our empirical analyses investigate the extent to which border-specific features and party cues shape voters' perspectives on cross-border cooperation and European integration. Our empirical findings show that not all political parties take similar positions on European integration and cross-border issues in border municipalities, whereas the citizenry is more consistent here: enhanced border integration is linked to more positive attitudes towards the EU. This, in turn, clearly illustrates that VAA data qualify as a valid source to produce new insights for research fields with data gaps.

2. Political Preferences Vis-À-Vis European Issues in Border Regions

For European democracies to function effectively, citizens expect their concerns and policy preferences to be represented by political parties and addressed by government institutions (Dahl, 1971; Esaiasson & Wlezién, 2017). However, although in European multi-level governance structures this is true for all levels of governance, with political parties competing about issues related to Europe from the local to the EU level (Braun & Schmitt, 2018), we are still very much focused on the national level when studying European democracy. In contrast, the specific needs of people at the local level often become diluted within larger governance frameworks, such as the regional and national level (Braun et al., 2025), and seem to play a minor role in local election campaigns (Braun & Hommel, 2025). This is particularly problematic for border regions. In these areas, home to around one third of EU citizens (European Commission, 2017, p. 2), the consequences of European integration have a direct impact on citizens' lives, as has recently been illustrated by border closures during the pandemic (Herbig et al., 2025) or by the implementation of stricter migration policies (Weber, 2022). Indeed, the European Commission (2021) identifies inner-European border regions as "living labs" for European integration, where the EU's functioning is tested daily. Also, scholars from border studies widely agree that border areas often have distinct characteristics (Banting et al., 2019) and their residents' specific preferences compared to those in non-border regions, reflecting shared interests such as cross-border mobility, healthcare, economic and political cooperation, and cultural exchange (Bürkner, 2020; Wassenberg & Reitel, 2020; Weber, 2022). Hence, patterns of party competition regarding the EU may be particular in these cases—both on the supply side (parties) and the demand side (citizens).

2.1. The Supply Side of Electoral Competition in European Border Regions

To date, almost no comparative research has looked into the supply side of party competition in border regions (but there are some case studies on specific regions, see e.g., Blatter, 2000; Braun et al., 2025; Lechevalier & Wielgoß, 2013). While this gap can be explained by the lack of data, it is an important limitation of the literature because political parties' attention to certain issues is strategic in nature and can play a key role in the formation of voters' attitudes (Adams et al., 2011; De Sio & Weber, 2014; Feddersen & Adams, 2022). Established theories of party competition as well as research on party competition about European issues in multi-level systems emphasise the relevance of both the saliency and the positioning of political parties toward certain issues (Adams et al., 2005; Green, 2007). Building on this, it seems plausible to expect that while the saliency of EU issues may be higher in elections to the European Parliament than on

other levels of government, party positions on EU issues should be similar across the different levels of government (Braun & Schmitt, 2018). Given that regional party branches reflect similar positions towards European integration as their national counterparts (Adendorf & Gross, 2025), we would assume that parties' positions on European integration at the local level should reflect the general stances of the respective parties.

However, it is less evident whether this general stance can be expected to extend to cross-border issues, given that case studies on the positioning of political parties in border regions did not yield conclusive findings (see also Bürkner, 2020). Although it is certainly true that the EU's vision of a Europe without borders has fostered the development of cooperation across borders, some studies also show negative repercussions of increasing cross-border integration (Durand et al., 2020; Wassenberg, 2010). In line with this, one could thus assume that even at the local level in border regions, Eurosceptic parties from the right follow the overall anti-European stance and oppose more cross-border integration, not only because it goes against the nationalistic ideology that underpins their political positions, but also because cross-border cooperation is not without its critical aspects. In Germany, the case researched in this article, this would lead us to expect that the AfD would oppose cross-border integration and favour policies such as strict border controls. Yet, such negative positions on cross-border issues could also backfire because they contrast with the daily experience of border residents, such as the ability to go shopping in neighbouring countries or having clients from across the border. Also, a strong regional identity linked to the cross-border aspects, for instance, due to a common history, such as in the German-French border region, may impede Eurosceptic parties from positioning themselves strongly against more cross-border integration. To inspect this relationship with our VAA data, we formulate the following first hypothesis:

H1. General party positions on EU integration correlate with the positions of local parties in border regions on cross-border issues.

2.2. The Demand Side of Electoral Competition in European Border Regions

In recent years, researchers have shown increased interest in studying the electoral geography of European integration across regions in Europe (Schraff & Pontusson, 2024) and, in particular, in border regions (for an early work, see Kuhn, 2011; Nasr & Rieger, 2024; Rehm et al., 2025). The specific interest of studying border regions is related to their position as areas where citizens experience the consequences of European integration in their daily lives. In border regions, citizens interact with residents from the neighbouring state(s), their language and culture; they experience the advantages and disadvantages of the single market (such as enhanced possibilities to purchase goods abroad, but restrictions when it comes to accessibility to health services); and they are immediately affected by restrictions to European integration, such as enhanced border controls. Therefore, the theoretical starting point of most of the existing studies goes back to transactional theory (Deutsch, 1953) and argues that both economic transactions (Sohn, 2014) as well as social interaction (Allport, 1954; Pettigrew, 1998) should make border residents more prone to develop more cosmopolitan attitudes and more positive attitudes toward their neighbours (Kuhn, 2015). However, recent studies that find a negative or no clear-cut relationship between border region residency and attitudes toward European integration have emphasised that the theoretical mechanism that links daily cross-border experiences to pro-European attitudes is more complicated. Indeed, in their recent study using panel data from Germany, Rehm et al. (2025, p. 503) “suggest moving beyond an overly optimistic idea of

border regions as ‘living labs of European integration’ to instead differentiate why some border regions display stronger European identification whilst others do not.”

Reflecting this need for a refined understanding, a very basic, but rarely tested assumption is whether citizens of border regions who are in favour of more cross-border cooperation also support European integration. While, on the one hand, transactional theory suggests that social interactions across borders can encourage more positive cosmopolitan attitudes in general (Mau, 2021; Taubenböck et al., 2023), regional identities do, on the other hand, not necessarily overlap with attitudes toward Europe in general. Whether or not a person has both a regional and a European identity at the same time depends on whether they are an exclusive or an inclusive regionalist (Brigevich, 2016; Praprotnik, 2020). Thus, it could well be that citizens who live in border regions favour increased border permeability to their direct neighbours but not European integration in general. Empirically, due to the lack of data on citizens’ opinions about cross-border issues, most studies could not investigate this question, but simply *assumed* that residency in border regions is related to more transnational experiences, leading to more cosmopolitan attitudes, which then translate into more positive views of European integration. However, with our VAA data collected at the local level in border regions, we are able to clarify whether this relationship actually holds because we can empirically measure how people position themselves to both cross-border issues and EU integration. While acknowledging the possibility of relationships going in both directions and given the strong role that transactional theory plays in the literature, we formulate our second hypothesis:

H2. Citizens who favour increased cross-border integration also support more European integration.

Studying this question, however, comes with the challenge of differentiating between European integration and cross-border integration as both may be driven by the same socio-economic variables. It seems plausible, for instance, that main correlates of Euroscepticism, such as education level, may also be related to attitudes toward more cross-border integration (Rehm et al., 2025). The same is true for political variables, such as left–right positions or political interest. Thus, if we want to study whether citizens who favour more cross-border integration also support more European integration, we need to account for such possible confounders. In fact, the specification of the theoretical model that needs to be tested empirically comes down to investigating a mediation effect. Therefore, as indicated in Figure 1, we expect that attitudes on

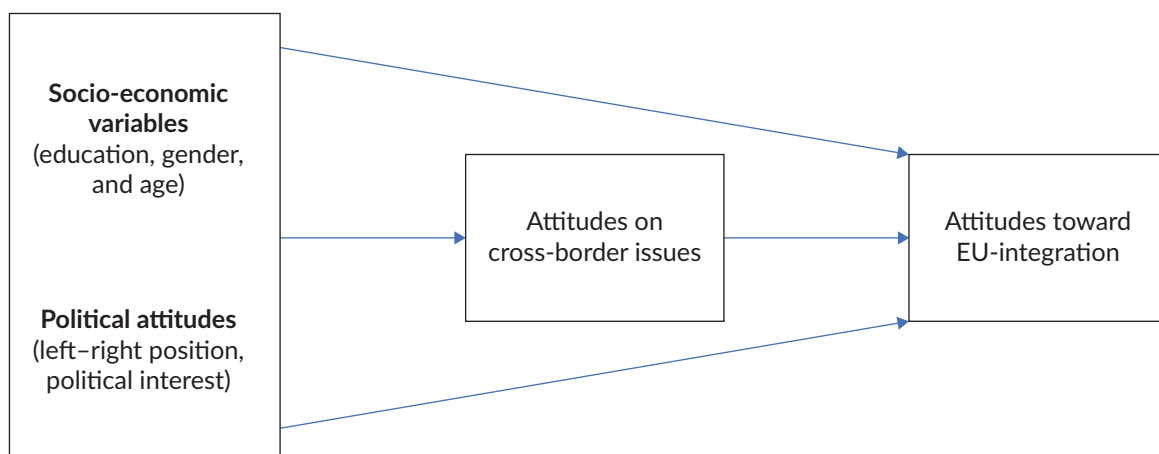


Figure 1. Simplified theoretical model.

cross-border issues mediate the effect of socio-economic and attitudinal variables that have been found to be correlated with EU preferences.

3. Research Design

3.1. Data and Case Selection

We analyse data from the VOTO VAA, which was conducted in 26 German municipalities in the run-up to the local elections held in June 2024, and realised by local teams of political scientists under the coordination of the non-profit enterprise VOTO gUG, which provides a digital platform for VAAs (for more information, see <https://www.voto.vote>). VAAs are digital tools that enable users to compare their personal views on political issues with the positions of political parties and derive a voting recommendation from this comparison (Garzia & Marschall, 2014, p. 377). Thus, they provide important insights into both the supply side of party competition (parties' positions on issues; see H1), as well as the demand side (attitudes, user preferences; see H2). As part of the VOTO project, the data was generated in a multi-stage process. In the first step, local teams of political scientists identified salient topics for their respective municipalities through media analysis, expert consultations, and analysis of electoral programs. Based on that, they developed specific statements relevant to their respective local politics, as well as bridge statements that they considered relevant across municipalities. Then, political parties were asked to rate the statements based on a five-point scale and to provide an explanation for their position in a free-text field. This resulted in a comparable database of party positions at the local level. In the next step, the statements and party positions were fed into an online tool, which was then made available to the public. Finally, the users themselves could rate the statements and receive voting advice based on their content-wise proximity to the parties. For academic purposes, we integrated a survey to collect the users' socio-demographic data directly after the voting advice. This resulted in a second database on the preferences and socio-demographic background of the VOTO users in different municipalities.

With respect to case selection, the analysis draws on VOTO data from five German municipalities (Saarbrücken, Merzig, Konstanz, Frankfurt (Oder), and Greifswald) in four German federal states (the regional political layer of the German federal system). They were selected due to their proximity to national borders, as depicted in Figure 2. Whereas Konstanz, Merzig, Saarbrücken, and Frankfurt (Oder) directly adjoin national borders, Greifswald is located slightly further inland, yet it still forms part of the wider border region with Poland. The Online Appendix A in the Supplementary File provides a short overview of these municipalities.

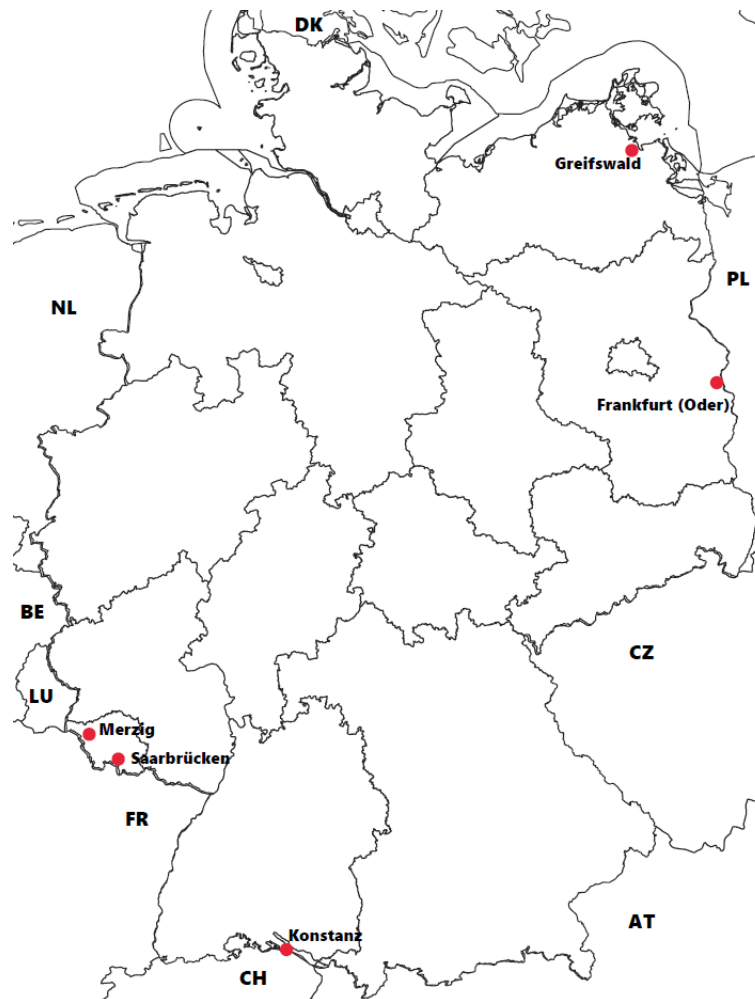


Figure 2. Map of the municipalities included in the analysis.

3.2. Operationalisation and Methods

To study party positions on European and border issues (H1), we use the statements shown in Figure 3. Although the number of EU and cross-border related items varies across municipalities, all VAAs contained two statements that directly relate to European integration, namely “{location} should campaign for more frequent random checks to be carried out near the border” and “The EU restricts the municipalities’ room for manoeuvre too much.” Additionally, for some of the municipalities, further statements on cross-border cooperation were included, such as on realising better cross-border train connections or creating a cross-border health system more accessible from different countries (see Figure 3). Both parties and users responded to these items on a five-point scale (*strongly disagree* to *strongly agree*). Methodologically, after presenting the descriptive findings on party competition in border regions, we compare party positions from VOTO VAA to general party positions on European integration, drawing on national-level data from the Chapel Hill Expert Survey (CHES), an expert survey that measures ideological positioning of political parties across Europe (Rovny et al., 2025).

	Saarbrücken	Merzig	Greifswald	Frankfurt (Oder)	Konstanz
{location} should campaign for more frequent random checks to be carried out near the border.	✓	✓	✓	✓	✓
The EU restricts the municipalities' room for manoeuvre too much.	✓	✓	✓	✓	✓
Seamless cross-border rail transport {between Germany, France and Luxembourg/between Germany and Poland} is more important than road construction/projects.	✓	✓	✗	✓	✗
{location} should invest more in a cross-border cycle network.	✓	✓	✗	✓	✗
In the daycare centres in {location}, the promotion of the {French/Polish} language should be intensified.	✓	✓	✗	✓	✗
In order to bring Europe closer to its citizens, more activities in the area of town twinning should be organised in {location}.	✓	✓	✓	✓	✗
Healthcare provision in {location} should be further expanded in order to create a cross-border healthcare corridor.	✓	✓	✗	✗	✗

Figure 3. Statements on European and border issues in VOTO (supply side/party data).

To examine the association between participants' views on border issues and EU integration (H2), we run multiple regression analyses including mediation. To do so, we use both citizens' positions on the VOTO border policy items (see Figure 3) as well as political attitudes and socio-demographics that participants gave in the post-VOTO survey. Importantly, only a subset of VOTO users, 4,407 users, participated in the survey. As descriptive analysis shows that certain groups in our VOTO VAA are significantly over- or underrepresented, we calculate a calibration weight based on the demographic characteristics of gender, age, and education in order to improve the representativeness of the sample. From a methodological perspective, iterative proportional fitting is employed to perform a ranking procedure that iteratively adjusts the weights for the three variables to align with the distribution in the target population (Sand & Kunz, 2020). Using the variables shown in Table 1, multiple regressions were all estimated with clustered standard errors (with municipalities as clusters) and included weights. For mediation analysis, we followed Baron and Kenny (1986) and estimated three models to see how coefficients change due to the inclusion of a mediator.

Table 1. Descriptive information on variables for demand side analyses (attitudinal data).

Variables	Obs	Mean	SD	Min	Max
EU integration attitudes	2,171	7.4	2.6	0.0	10.0
Border controls (higher: not agree)	21,764	52.1	37.2	0.0	100.0
EU restricts municipalities (higher: not agree)	21,763	54.5	30.8	0.0	100.0
Border policy factor	8,679	0.0	1.0	−2.8	1.5
Gender (1 - female)	3,984	0.47	0.5	0.0	1.0
Age	3,959	44.7	16.5	16.0	100.0
Education	3,853	5.4	1.0	1.0	6.0
Political interest (higher: more interested)	3,812	3.7	0.9	1.0	5.0
Left–right position (higher: more right)	3,295	4.5	1.9	1.0	11.0
Satisfaction with government (higher: more satisfied)	3,145	−0.9	2.8	−5.0	5.0

Notes: Border controls and the extent to which the EU is perceived to restrict municipalities are measured on a five-point scale (0, 25, 50, 75, 100), with higher values indicating lower support for border controls and for the statement that the EU restricts municipalities. Political interest is also measured on a five-point scale, with higher values representing greater political interest, while left–right orientation and satisfaction with the government are measured on an 11-point scale, EU integration on a scale from 0 to 10, and education on an ordinal scale with six categories (for the ranges, see the Min–Max columns in the table). Age runs from 16 to 100 years, and the border policy factors range from −2.8 to 1.5 as generated by a factor analysis. The border policy factor was generated via a factor analysis based on participants' attitudes regarding four items of the VOTO: town twinnings, language of the neighbour, more cross-border train tracks, and more cross-border bike lines, all measured on a five-point scale (0, 25, 50, 75, 100). Higher values indicate higher support.

4. Empirical Analyses

4.1. The Supply Side: Party Competition on European and Border Issues in Border Regions

In this section, we provide a descriptive analysis of party positions on questions of European and cross-border integration in the selected border municipalities. Figure 4 illustrates the specific responses of the main parties, broken down by municipality.

In the first step of the analysis, we zoom in on how the parties rate single aspects of European integration in the VOTO VAA, namely the extension of border controls and the restriction of their room for manoeuvre by the EU. We also analyse how they locally differ in their response behaviour regarding each of the aspects presented below:

- *Border controls:* The examination of parties' positions on border controls reveals a heterogeneous landscape. The AfD adopts a strongly affirmative position, advocating unequivocally for strengthening border controls (minimum = 75, 1st quartile = 87.5, median = 100, 3rd quartile = 100, maximum = 100). It is striking that the CDU is strongly divided on the issue; however, on average, local CDU sections indicate that border controls need to be strengthened (minimum = 0, 1st quartile = 50, median = 75, 3rd quartile = 100, maximum = 100). Only in Saarbrücken, the party strictly opposes, while in Konstanz and Frankfurt (Oder), it strongly supports border controls. Equally, the FDP maintains an undecided stance, showing a strong variance across municipalities (minimum = 0, 1st quartile = 0, median = 50, 3rd quartile = 50, maximum = 75). Bündnis 90/Die Grünen expresses clear opposition to the implementation of border controls, except for the two Saarland municipalities

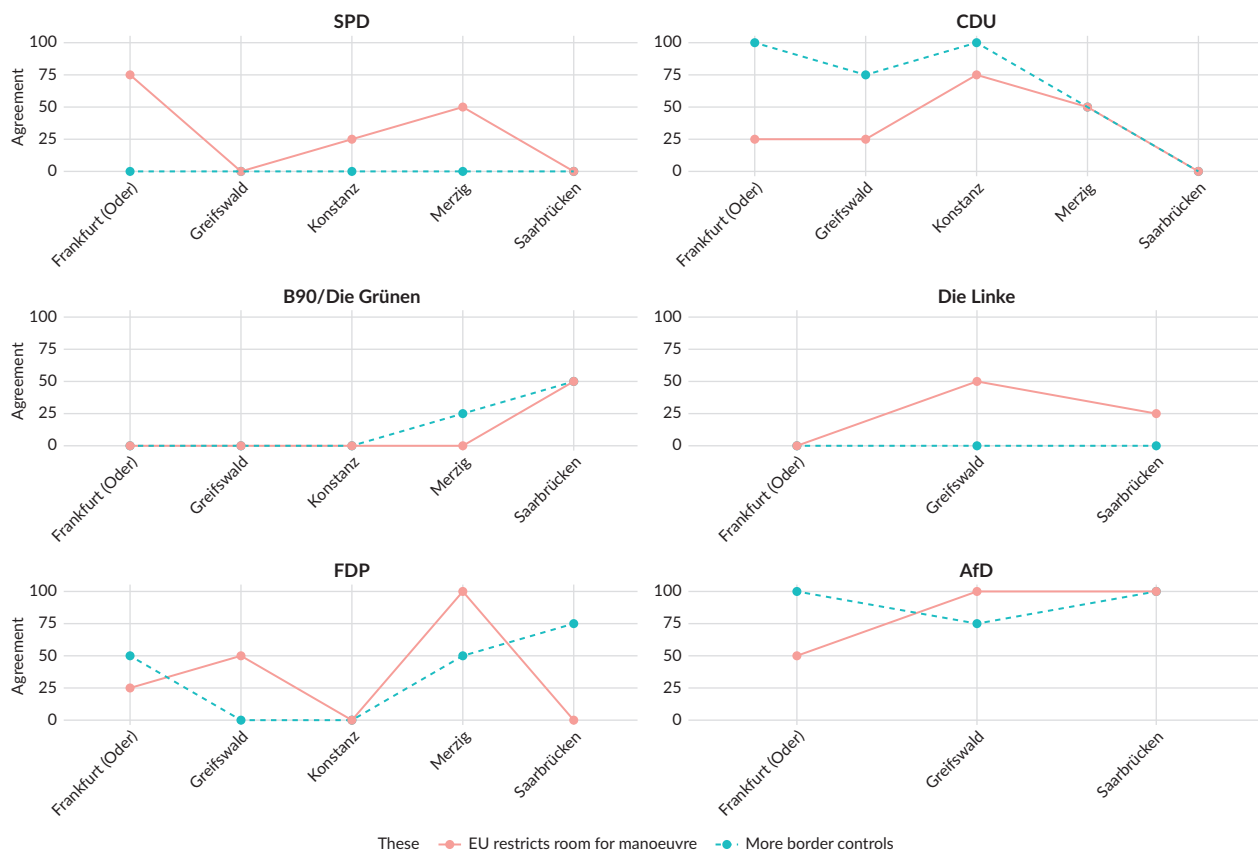


Figure 4. Party positions in the selected municipalities. Source: Own calculation based on VOTO VAA data (Stecker et al., 2025).

(minimum = 0, 1st quartile = 0, median = 0, 3rd quartile = 0, maximum = 25). Finally, both SPD (minimum = 0, 1st quartile = 0, median = 0, 3rd quartile = 0, maximum = 0) and Die Linke (minimum = 0, 1st quartile = 0, median = 0, 3rd quartile = 0, maximum = 0) respond coherently across border municipalities, positioning themselves unanimously against border controls.

- *EU restricts room for manoeuvre*: Moving on to the next question of how far the EU restricts the municipality's room for manoeuvre, we observe a maximum variation within the FDP (minimum = 0, 1st quartile = 0, median = 25, 3rd quartile = 50, maximum = 100). While the FDP in Merzig fully perceives the EU as a restriction on the municipalities, it positions itself at the other extreme in Konstanz and Saarbrücken. Overall, the FDP tends not to consider that the EU narrows its room for manoeuvre. Slightly less than FDP, but also divided on this question, the SPD, in sum, equally tends to the view that the EU does not restrict municipalities' freedom of action (minimum = 0, 1st quartile = 0, median = 25, 3rd quartile = 50, maximum = 75). The latter also applies to both the CDU (minimum = 0, 1st quartile = 25, median = 25, 3rd quartile = 50, maximum = 75), with the CDU in Saarbrücken as pro-European and the CDU in Konstanz less clearly pro-European, and Die Linke, which shows less variation (minimum = 0, 1st quartile = 12.5, median = 25, 3rd quartile = 37.5, maximum = 50). Bündnis 90/Die Grünen are unanimously not considering the EU to restrict them, with the exception of Saarbrücken (minimum = 0, 1st quartile = 0, median = 0, 3rd quartile = 25, maximum = 50). In stark contrast, the AfD is united in its opposition to the EU, with the outlier of Frankfurt (Oder) (minimum = 50, 1st quartile = 75, median = 100, 3rd quartile = 100, maximum = 100).

- Mapping local party position clusters:** In the next step, we plot party positions on both issues to identify parties or locations with similar response patterns. Thereby, as depicted in Figure 5, five party clusters can be identified. Firstly, a large pro-European cluster strongly supports the free movement of goods and people, and does not perceive the EU as any restriction on its ability to act. This group includes all local groups of Die Linke; all branches of Bündnis 90/Die Grünen, except Saarbrücken; the SPD, with the exception of Frankfurt (Oder); as well as several local lists. The CDU in Saarbrücken and the FDP in Konstanz constitute local outliers within this pro-European group. Two further clusters hold more ambivalent views on questions of European integration: a small cluster on the one hand believes that the EU restricts its freedom to act, and on the other hand strictly opposes the extension of border controls, including the SPD in Frankfurt (Oder), the FDP in Merzig, and JFK in Konstanz. A third cluster wants to enforce border controls, but does feel restricted by the EU. This is the second-largest cluster, comprising predominantly FDP and CDU local groups. What remains are two Eurosceptic clusters: One cluster would like to strengthen border controls and believes that the EU is restricting its room for manoeuvre. This group includes the Frankfurt (Oder) branch of the AfD, the Freie Wähler, and the CDU in Konstanz. Finally, a fifth cluster is severely Eurosceptic, perceiving itself severely restricted by the EU and strictly calling for more border controls. This includes the AfD local branches and local parties in Greifswald. We can thus summarise that, according to our VOTO data, CDU, FDP, and AfD tend to favour cross-border cooperation over European integration, whereas the degree of support for cross-border cooperation and European integration is comparable for both the SPD and Bündnis 90/Die Grünen.

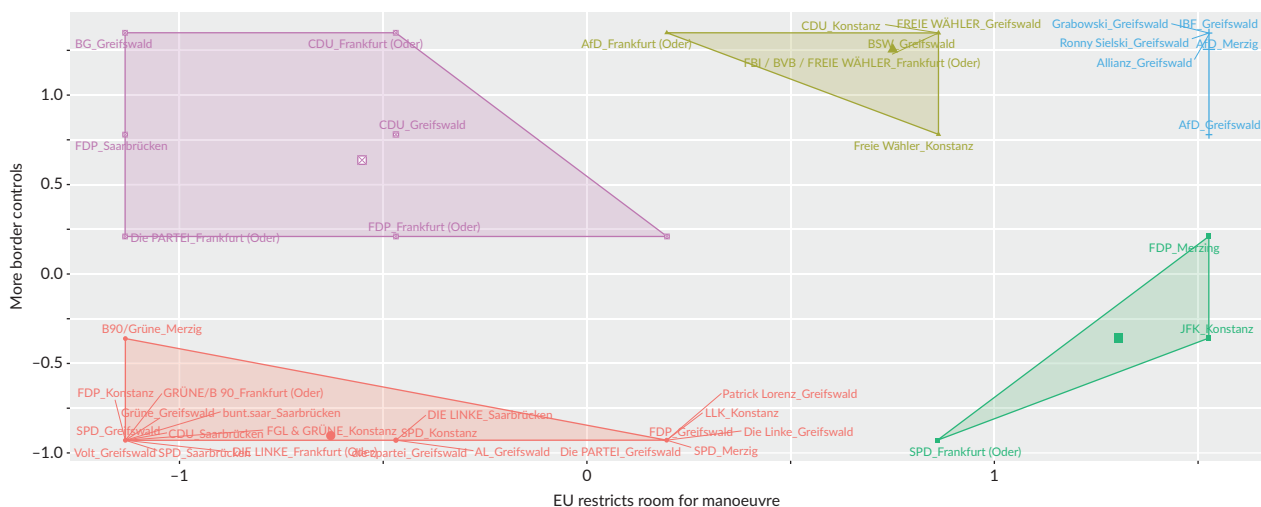


Figure 5. Clustered party positions. Source: Own calculation based on k-means clustering, data from the VOTO VAA (Stecker et al., 2025). Note: Higher values reflect stronger agreement with the items.

In the second step, we zoom out to test our first hypothesis, which assumes that party competition in border municipalities corresponds to general party positions. Figure 6 shows a comparison of party positions at the national level (general position on European integration, as derived from CHES data) with two local-level VOTO variables: first, *party positions on European integration*—mean value of two European issues: (a) border controls and (b) restriction of the municipalities' room for manoeuvre by the EU—and, second, *party positions on cross-border integration*—mean value of several border issues, use of items where surveyed: (a) cross-border rail transport, (b) teaching of neighbour language, (c) cross-border cycle network, (d) cross-border healthcare, and (e) town twinning.

For some parties, we observe significant discrepancies between party positions at the local and the national level. While Bündnis 90/Die Grünen, SPD, and AfD largely adhere to their national party positions, this is not the case for Die Linke and CDU, and only to a limited extent for the FDP. It is striking that Die Linke in the border municipalities is much more pro-European than at the national level, where the party takes a rather Eurosceptic stance. This may have to do with the fact that members of the party in border regions can feel the direct implications of cross-border cooperation much more strongly in concrete projects whereas the national party program is more on high-politics issues and wants to signal to voters that Die Linke is sceptical when it comes to liberal economic policies or the increased focus on military and defence issues at the EU level. In comparison, for the AfD, also a Eurosceptic party, this pattern seems less pronounced, probably because the general anti-elitist discourse, which also caters to protest voters, looms too large, and the party cannot deviate from the overall stance in border regions. Conversely, the CDU is significantly less pro-European in border municipalities (and especially so in Konstanz) than at the national level, with the exception of Saarbrücken. At least for Konstanz, a possible explanation may be the fact that the municipality's border is with Switzerland—a non-EU country—where at least one of our items used to measure a party's stance toward European integration (border controls) comes with a special note, as Switzerland has never stopped controlling its border.

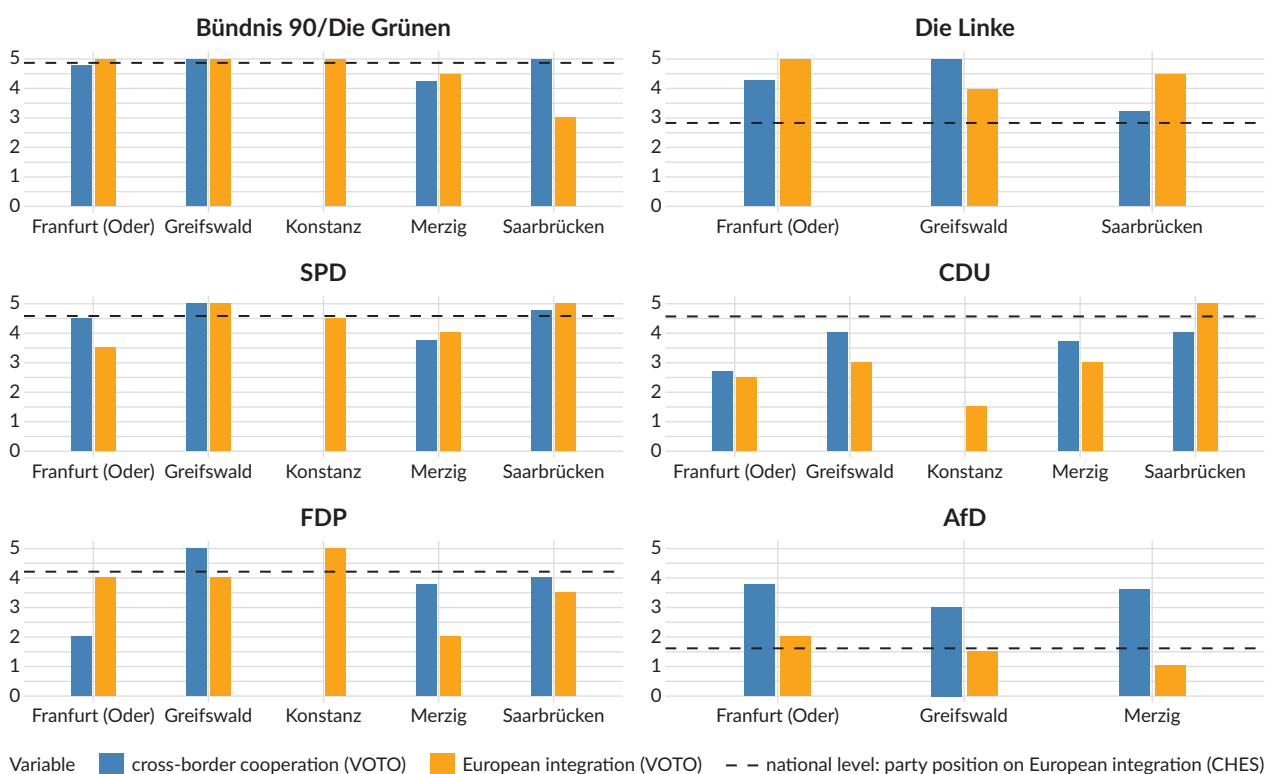


Figure 6. Comparison of party positions in border municipalities and at the national level. Notes: The black line shows national party positions on European integration according to CHES data (Rovny et al., 2025), blue bars represent the mean value for items on border issues based on local VOTO data (Stecker et al., 2025), and orange bars represent the mean value for items on European issues based on local VOTO data, with higher values indicating a more positive view of European and cross-border issues. The illustration includes established parties that exist at both the federal and local levels; lists that only exist at the local level are not included here.

In general, it is important to keep in mind that these results may also reflect mainly the political positions of a party's local branches (and members), which may also be idiosyncratic to a certain extent. Keeping in mind these limitations and given the ambiguous results, we cannot accept H1 without qualification.

4.2. The Demand Side: Voters' Positions on European and Border Issues

Having analysed the supply side, namely party competition on European and border issues, we now turn to the individual level and present results from voters' positions. As discussed in Section 3.1, we draw on two different data sources: First, we use data from the VOTO itself, namely the positions of voters in the municipalities on border issues and European issues. Second, we leverage data from a survey that followed the VOTO (with a restricted sample size) and included attitudinal questions and socio-demographic variables. This allows us to test our second hypothesis and to investigate to what extent attitudes toward border issues are correlated with opinions about European integration.

4.2.1. Correlations of Positions on Border Issues and European Issues as Measured in the VOTO

To give an overview of how citizens who participated in the VOTO positioned themselves toward EU issues and the border region, we have calculated simple Pearson intercorrelations between border-related and EU-related issues (see Figure 3 for wording; the coefficients are to be found in Appendix D in the Supplementary File). The correlations indicate that, first, the different border issues (e.g., attitudes on border trains and border bike lines) are moderately to strongly intercorrelated (those who want more cross-border train tracks also want more cross-border bike lines). The same is true for EU-related issues (those who want more border controls also think that the EU restricts municipalities). In contrast, we only observe a weak correlation between the border issues and the two EU-related variables, although we see no negative associations.

4.2.2. Regression Analysis Based on VOTO and Survey Data: Cross-Border Integration Equals European Integration?

While the correlational patterns already give some slight indication that a positive attitude toward cross-border cooperation may go hand in hand with a preference for European integration, the bivariate associations could also be driven by confounding variables linked to more general political attitudes or socio-demographic characteristics, such as education. In order to account for these possible confounders, we have estimated several multiple regressions using the data collected from the short survey that accompanied the VAA, which includes information on socio-demographics and political attitudes. Combining survey and VOTO data allows us to test our H2, namely, whether citizens who favour increased cross-border integration also support more European integration. All regressions presented in Table 2 were estimated using clustered standard errors (for the municipalities) and weighted for socio-demographics (Appendix B in the Supplementary File includes the same models using non-weighted data). The analyses do not cover Konstanz as no border-specific questions were asked there (see Figure 3). As dependent variable, we used three different measures of attitudes toward the EU: (a) A 10-point scale based on a question of the European Social Survey asking whether EU integration should be further extended or not; and the two variables from the VOTO, namely, (b) citizens' attitudes toward border controls, and (c) their position toward the statement that the EU restricts a municipalities scope of action. As independent variables, we used (a) a factor based on the four items on border policies included above (on transportation, language, and town

twinings) and (b) the town twinings variable alone to increase the number of respondents by including the VOTO in Greifswald.

Table 2. Regression analysis: Border integration and European integration (weighted data).

	(1) EU integration attitudes	(2) Border controls	(3) EU restricts the municipality	(4) EU integration attitudes	(5) Border controls	(6) EU restricts the municipality
Border policy factor	0.88* (8.36)	5.53** (13.79)	6.20* (8.75)			
Twin towns				0.027** (8.71)	−0.0029 (−0.04)	0.19 (2.63)
Gender (Ref: Man)	−0.018 (−0.07)	4.98 (1.84)	−1.89 (−0.80)	−0.0020 (−0.01)	1.34 (0.45)	2.82 (1.38)
Age	−0.012 (−4.31)	−0.14 (−4.09)	−0.051 (−0.81)	−0.0015 (−0.44)	−0.15* (−3.50)	−0.23 (−1.24)
Education	0.26 (3.01)	5.64* (7.37)	3.94** (24.01)	0.24** (5.87)	6.31** (9.23)	5.18* (3.66)
Political interest	0.38* (7.67)	−0.094 (−0.11)	7.79 (3.82)	0.22 (2.67)	3.29 (1.28)	3.97 (1.49)
Left-right	−0.17 (−2.45)	−4.31 (−1.59)	−1.76 (−3.25)	−0.31** (−5.75)	−5.68* (−3.15)	−2.06*** (−11.78)
Satisfaction government	0.20* (8.67)	2.81 (1.58)	2.37 (3.98)	0.25** (6.83)	2.70 (2.49)	3.13*** (11.65)
Constant	6.10** (22.78)	53.4** (15.52)	22.5** (15.06)	5.07** (7.42)	45.1*** (22.41)	23.2 (1.91)
N	974	1,034	1,034	1,383	1,474	1,474
R ²	0.363	0.256	0.303	0.345	0.284	0.297
weights	yes	yes	yes	yes	yes	yes

Notes: *t*-statistics in parentheses, * $p < .1$, ** $p < .05$, *** $p < .01$. All regressions with clustered standard errors (clusters are the municipalities).

The results from the regression analyses are rather straightforward: Independent of the measurement of the independent variables used, citizens who prefer more cross-border integration also exhibit more positive attitudes toward the EU and think that the EU does not restrict municipalities' scope of action too severely. While the respective coefficients are highly significant at conventional levels, the relationship with the border control item is not (while still pointing in the expected direction). Hence, controlled for possible confounders, these results support our expectation formulated in H2 according to which an enhanced border integration is linked to more positive attitudes toward the EU. The control variables included in the estimations perform as expected: more politically interested, more educated, and ideologically more left-leaning participants also support EU integration more strongly. The same is true for those who are more satisfied with the government. However, no clear-cut findings emerge for gender and age. In addition to the direct effects uncovered in the regression analysis, which indicate that participants' preferences on cross-border integration are linked with their attitudes toward European integration, a further interrogation concerns the nature of the relationship between those two variables and political attitudes included as

control variables in the regression analysis. In fact, as theorised in Section 2.1 (see Figure 1), it could well be that part of the influence of political attitudes and socio-demographics that have been identified as drivers of Eurosceptic views in the literature is mediated by citizens' preferences on cross-border issues. A full mediation would come down, for instance, to the fact that highly educated people living in border regions only seem to favour more European integration, but mainly want more cross-border cooperation, which is in turn related to positive views about Europe. In this case, cross-border integration would be the key variable in that equation and fully soak up the direct effect between education level and attitudes toward Europe (see our theoretical considerations in Section 2.1).

We have investigated this question in a mediation analysis displayed in Table 3 (only for Frankfurt (Oder), Merzig, and Saarbrücken). The first model shows the direct relationships between our attitudinal variables and support for more EU integration, whereas the second model shows the associations of the same variables with support for cross-border integration, our mediator. Finally, Model 3 displays the full model including the border factors as predictors for European integration attitudes. Following Baron and Kenny (1986), we start by analysing the direct relationship between political attitudes and socio-demographics and EU integration attitudes (Model 1). The results show that more politically interested and more left-leaning persons hold more EU-positive opinions. Interestingly, though, turning to model 2, these variables are not significantly related to the border policy factor that takes up preferences for more cross-border integration. Hence, a pre-requisite for mediation does not seem to be fulfilled. Finally, the third step also indicates that the opinions on border issues exhibit an independent effect on attitudes toward the EU: while political interest and a left position remain

Table 3. Mediation analysis.

	(1) EU integration attitudes	(2) Border policy factor	(3) EU integration attitudes
Gender (Ref: Man)	−0.089 (−0.27)	−0.079 (−0.71)	−0.018 (−0.07)
Age	−0.012 (−2.36)	0.00019 (0.09)	−0.012 (−4.31)
Education	0.27 (2.02)	0.0058 (0.13)	0.26 (3.01)
Political interest	0.50** (16.42)	0.19 (3.79)	0.38* (7.67)
Left-right	−0.29* (−11.45)	−0.13 (−1.73)	−0.17 (−2.45)
Satisfaction government	0.24 (5.26)	0.056 (2.76)	0.20* (8.67)
Border policy factor			0.88* (8.36)
_cons	6.41** (63.40)	0.074 (0.12)	6.10** (22.78)
N	974	1034	974
R ²	0.280	0.178	0.363
Weight	yes	yes	yes

Notes: *t*-statistics in parentheses, * $p < .1$, ** $p < .05$, *** $p < .01$.

significantly associated with attitudes toward EU integration, the views about border policies are an additional variable significantly associated with EU attitudes—a point that is also reflected in the increase of R^2 compared to Model 1. In sum, we can conclude that independently of well-known correlates of Euroscepticism, VOTO participants who favour more cross-border cooperation also tend to want closer EU integration.

5. Conclusions

Due to data restrictions, studies of European democracy still predominantly focus on relationships between the national and the EU level. Studying the field of party competition within the EU's multi-level system, this article highlights the role of European border regions and shows how dynamics in these inner-EU regions provide valuable insights into research questions about European democracy. Despite being home to around one-third of EU citizens, little is known about patterns of party competition, as well as citizens' specific demands and preferences in border regions. We were finally able to address this issue by using innovative data from the VOTO VAA, which was specifically designed for local elections in Germany in 2024. Investigating how European democracy functions at a local level by examining the positions of political parties and the political preferences of citizens on four German border regions—two on the eastern border and two on the western border—our study shows the extent to which border-specific features and party cues influence voters' views on cross-border cooperation and European integration.

Our empirical findings reveal interesting insights into both the supply and demand sides of political competition. Regarding the supply side, the empirical analysis only partially confirmed our initial hypothesis that party competition in border municipalities corresponds to general party positions. This is indeed only the case for certain parties. Both CDU and Die Linke in border municipalities significantly deviate from their respective national-level party position. Notably, we observe considerable variation within a party across border municipalities (e.g., within the FDP). Although the analysis reveals many local outliers, we cannot claim that parties in a given municipality consistently support or oppose European issues. While in Saarbrücken, the CDU is outstandingly pro-European, Bündnis 90/Die Grünen can be considered more Eurosceptic in comparison to the other very pro-European branches of the Greens in other border municipalities. For the demand side of political competition, our findings show that EU preferences correlate with attitudes on cross-border cooperation. Altogether, we can thus conclude that not all political parties take similar positions on European integration and cross-border issues in border municipalities, whereas the citizenry is more consistent here: enhanced border integration is linked to more positive attitudes towards the EU. This, in turn, clearly illustrates that VAA data qualifies as a valid data source to produce new insights for research fields with data gaps.

Based on these findings, we can identify at least three important avenues for further research: First, future studies should take into account the characteristics of different types of (European) border-regions. While our limited sample and the multiple dimensions of differences between the towns do not allow us to draw systematic conclusions, an interesting avenue could be to see how citizens experience the border itself. For example, the German-Polish border experienced border controls during the field period of VOTO, while the German-French border did so only to a limited degree. Second, in our study, we were only able to study positional preferences, but—due to data restrictions—did not address the salience. Given that different party families might have different incentives to emphasise cross-border integration, this perspective could also be insightful. The situation in terms of incentive structures seems to be unclear for political parties—i.e., is it

strategically useful or not for political parties to put emphasis on cross-border issues? In line with theoretical work on party competition, we can act on the assumption that such an environment typically leads to a de-emphasis of such issues, and this is actually what has been shown by a case study in the German-French border region (Braun & Hommel, 2025). Third, future studies should definitely continue to use innovative data sources, such as VAA data, to study research questions with an interest in local- or regional-level problems to overcome the issue of a national bias in studies of European democracy.

Acknowledgments

The authors would like to thank the students at Saarland University and the many research assistants participating in the elaboration and fielding of VOTO as well as the VOTO core group based at TU Darmstadt for their valuable support. Moreover, the authors are grateful to the anonymous reviewers for their comments and suggestions.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability

The full data of the VOTO-project can be found at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PAH9O1>

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

- Adams, J., Ezrow, L., & Somer-Topcu, Z. (2011). Is anybody listening? Evidence that voters do not respond to European parties' policy statements during elections. *American Journal of Political Science*, 55(2), 370–382. <https://doi.org/10.1111/j.1540-5907.2010.00489.x>
- Adams, J. F., Merrill, S., & Grofman, B. (2005). *A unified theory of party competition: A cross-national analysis integrating spatial and behavioral factors*. Cambridge University Press.
- Adendorf, A., & Gross, M. (2025). Under pressure? How regional dominant parties shift their EU positions when challenged by Euroskeptic parties. *Journal of European Integration*, 47(7), 1081–1105. <https://doi.org/10.1080/07036337.2024.2432326>
- Allport, G. W. (1954). *The nature of prejudice* (3rd ed.). Addison-Wesley.
- Banting, K., Nagel, J., Schafer, C., & Westlake, D. (2019). Assessing performance. National versus regional patterns. In P. J. Quirk (Ed.), *The North American experiment: Institutions and policymaking in Canada and the United States* (pp. 290–319). Oxford University Press.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Blatter, J. (2000). Emerging cross-border regions as a step towards sustainable development? Experiences and Considerations from examples in Europe and North America. *International Journal of Economic Development*, 2(3), 402–439. <https://ssrn.com/abstract=3009670>
- Bornschier, S. (2010). The new cultural divide and the two-dimensional political space in Western Europe. *West European Politics*, 33(3), 419–444. <https://doi.org/10.1080/01402381003654387>

- Braun, D., & Hommel, F. (2025). Euroskeptizismus in der deutsch-französischen Grenzregion. Eine Tageszeitungsanalyse zu den Kommunalwahlkämpfen in Saarbrücken und Forbach. In D. Braun, U. Connor & K. Höfer (Eds.), *Grenzregionen als Laboratorien der Europaforschung* (pp. 291–312). Nomos. <https://doi.org/10.5771/9783748954606-291>
- Braun, D., & Schmitt, H. (2018). Different emphases, same positions? The election manifestos of political parties in the EU multilevel electoral system compared. *Party Politics*, 26(5), 640–650. <https://doi.org/10.1177/1354068818805248>
- Braun, D., Carteny, G., Hartland A., Jabra, D., & Navarrete, R. M. (2025). Political parties' policy preferences on environmental issues in German border regions. In D. Braun, U. Connor & K. Höfer (Eds.), *Grenzregionen als Laboratorien der Europaforschung* (pp. 251–277). Nomos. <https://doi.org/10.5771/9783748954606-251>
- Brigevich, A. (2016). Regional identity and support for Europe: Distinguishing between cultural and political social identity in France. *Regional & Federal Studies*, 26(4), 475–507. <https://doi.org/10.1080/13597566.2016.1223057>
- Bright, J., Garzia, D., Lacey, J., & Trechsel, A. (2016). Europe's voting space and the problem of second-order elections: A transnational proposal. *European Union Politics*, 17(1), 184–198. <https://doi.org/10.1177/1465116515588965>
- Bürkner, H.-J. (2020). Europeanisation versus Euroscepticism: Do borders matter? *Geopolitics*, 25(3), 545–566. <https://doi.org/10.1080/14650045.2020.1723964>
- Dahl, R. A. (1971). *Polyarchy. Participation and opposition*. Yale University Press.
- De Sio, L., & Weber, T. (2014). Issue yield: A model of party strategy in multidimensional space. *American Political Science Review*, 108(4), 870–885. <https://doi.org/10.1017/S0003055414000379>
- De Vries, C. E., & Hobolt, S. B. (2020). *Political entrepreneurs. The rise of challenger parties in Europe*. Princeton University Press.
- Debus, M., & Schulte, F. (2024). How party competition shapes ethnic parties' positions on migration and immigration. *Party Politics*, 30(1), 126–138. <https://doi.org/10.1177/13540688221136107>
- Deutsch, K. W. (1953). The growth of nations: Some recurrent patterns of political and social integration. *World Politics*, 5(2), 168–195. <https://doi.org/10.2307/2008980>
- Dinas, E., Trechsel, A. H., & Vassil, K. (2014). A look in the mirror: Preferences, representation and electoral participation. *Electoral Studies*, 36, 290–297. <http://dx.doi.org/10.1016/j.electstud.2014.04.011>
- Durand, F., Decoville, A., & Knippschild, R. (2020). Everything all right at the internal EU borders? The ambivalent effects of cross-border integration and the rise of Euroscepticism. *Geopolitics*, 25(3), 587–608. <https://doi.org/10.1080/14650045.2017.1382475>
- Esaiasson, P., & Wlezien, C. (2017). Advances in the study of democratic responsiveness: An introduction. *Comparative Political Studies*, 50(6), 699–710. <https://doi.org/10.1177/0010414016633226>
- European Commission. (2017). *Boosting growth and cohesion in EU border regions. Communication from the Commission to the Council and the European Parliament* (COM(2017) 534 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017DC0534>
- European Commission. (2021). *EU Border Regions: Living labs of European integration*. https://ec.europa.eu/regional_policy/information-sources/publications/reports/2021/eu-border-regions-living-labs-of-european-integration_en
- Feddersen, A., & Adams, J. F. (2022). Public opinion backlash in response to party messages: A case for party press releases. *Electoral Studies*, 80, Article 102532. <https://doi.org/10.1016/j.electstud.2022.102532>
- Fivaz, J., Louwerse, T., & Schwarz, D. (2014). Keeping promises: Voting advice applications and political representation. In D. Garzia & S. Marschall (Eds.), *Matching voters with parties and candidates. Voting advice applications in comparative perspective* (pp. 197–215). ECPR Press.

- Garzia, D., & Marschall, S. (Eds.). (2014). *Matching voters with parties and candidates. Voting advice applications in comparative perspective*. ECPR Press.
- Green, J. (2007). When voters and parties agree: Valence issues and party competition. *Political Studies*, 55(3), 629–655. <https://doi.org/10.1111/j.1467-9248.2007.00671.x>
- Gross, M. (2022). Does anyone care? Cohesion policy issues in sub-national politics. *JCMS: Journal of Common Market Studies*, 60(6), 1538–1555. <https://doi.org/10.1111/jcms.13307>
- Gross, M., & Debus, M. (2018). Does EU regional policy increase parties' support for European integration? *West European Politics*, 41(3), 594–614. <https://doi.org/10.1080/01402382.2017.1395249>
- Gross, M., & Jankowski, M. (2020). Dimensions of political conflict and party positions in multi-level democracies: Evidence from the Local Manifesto Project. *West European Politics*, 43(1), 74–101. <https://doi.org/10.1080/01402382.2019.1602816>
- Herbig, L., Unan, A., Kuhn, T., Rodríguez, I., Rodon, T., & Klüver, H. (2025). Closed borders, closed minds? Covid-related border closures, EU support and hostility towards immigrants. *European Journal of Political Research*, 64(4), 1923–1944. <https://doi.org/10.1111/1475-6765.70025>
- Hooghe, L., & Marks, G. (2018). Cleavage theory meets Europe's crises: Lipset, Rokkan, and the transnational cleavage. *Journal of European Public Policy*, 25(1), 109–135. <https://doi.org/10.1080/13501763.2017.1310279>
- Hutter, S., Grande, E., & Kriesi, H. (2016). *Politicising Europe: Integration and mass politics*. Cambridge University Press.
- Kuhn, T. (2011). Europa ante portas: Border residence, transnational interaction and Euroscepticism in Germany and France. *European Union Politics*, 13(1), 94–117. <https://doi.org/10.1177/1465116511418016>
- Kuhn, T. (2015). *Experiencing European integration: Transnational lives and European identity*. Oxford University Press.
- Lechevalier, A., & Wielgoths, J. (2013). *Borders and border regions in Europe. Changes, challenges and chances*. Transcript. <https://doi.org/10.14361/transcript.9783839424421>
- Mau, S. (2021). *Sortiermaschinen: Die Neuerfindung der Grenze im 21. Jahrhundert*. CH Beck.
- Nasr, M., & Rieger, P. (2024). Bringing geography back in: Borderlands and public support for the European Union. *European Journal of Political Research*, 63(4), 1328–1350. <https://doi.org/10.1111/1475-6765.12652>
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, 49, 65–85. <https://doi.org/10.1146/annurev.psych.49.1.65>
- Praprotnik, K. (2020). Regional Europeans. The relationship between social identities and EU support in Austria. In E. Donat, S. Meyer & G. Abels (Eds.), *European regions. perspectives, trends and developments in the 21st century* (pp. 43–58). Transcript. <https://doi.org/10.14361/9783839450697>
- Rehm, M., Schröder, M., & Wenzelburger, G. (2025). Border regions as nuclei of European integration? Evidence from Germany. *JCMS: Journal of Common Market Studies*, 63(2), 485–506. <https://doi.org/10.1111/jcms.13607>
- Röth, L., Saldivia Gonzatti, D., Kaftan, L., & Kaiser, A. (2025). Studying multi-level systems with cross-level data: Introducing three integrated datasets. *British Journal of Political Science*, 55, Article e19. <https://doi.org/10.1017/S0007123424000553>
- Rovny, J., Bakker, R., Hooghe, L., Jolly, S., Marks, G., Polk, J., Steenbergen, M., & Vachudova, M. (2025). *25 years of political party positions in Europe: The Chapel Hill expert survey, 1999–2024* (Working paper). <https://www.chesdata.eu/2024-chapel-hill-expert-survey-ches>

- Sand, M., & Kunz, T. (2020). *Gewichtung in der Praxis*. GESIS—Leibniz Institute for the Social Sciences. https://doi.org/10.15465/gesis-sg_030
- Scharpf, F. W. (2019). Multilevel democracy: A comparative perspective. In N. Behnke, J. Broschek, & J. Sonnicksen (Eds.), *Configurations, dynamics and mechanisms of multilevel governance* (pp. 249–271). Springer. https://doi.org/10.1007/978-3-030-05511-0_14
- Schraff, D., & Pontusson, J. (2024). Falling behind whom? Economic geographies of right-wing populism in Europe. *Journal of European Public Policy*, 31(6), 1591–1619. <https://doi.org/10.1080/13501763.2023.2278647>
- Sohn, C. (2014). Modelling cross-border integration: The role of borders as a resource. *Geopolitics*, 19(3), 587–608. <https://doi.org/10.1080/14650045.2014.913029>
- Stecker, C., Dieing, T. I., Oblong, J., Strobel, S., Braun, D., Deiss-Helbig, E., Gessler, T., Hein, M., Höhne, B., Kueppers, A., Linhart, E., Müller, J., Pravemann, S., Reiser, M., Steiner, N. D., Thomeczek, J.P., Völkl, K., Weisskircher, M., Wenzelburger, G., & Zissel, P. (2025). *Voting advice applications for local elections (Version V1)* [Data set]. Harvard Dataverse. <https://doi.org/10.7910/DVN/PAH9O1>
- Taubenböck, H., Otto, C., Gülzau, F., & Mau, S. (2023). Border regions across the globe: Analyzing border typologies, economic and political disparities, and development dynamics. *Applied Geography*, 151, Article 102866. <https://doi.org/10.1016/j.apgeog.2022.102866>
- Thomas, A., & Jones, E. (2025). The impact of ‘populism’ on European public policy. *Journal of European Public Policy*, 32(5), 1061–1072. <https://doi.org/10.1080/13501763.2025.2468901>
- Treib, O. (2021). Euroscepticism is here to stay: What cleavage theory can teach us about the 2019 European Parliament elections. *Journal of European Public Policy*, 28(2), 174–189. <https://doi.org/10.1080/13501763.2020.1737881>
- Vasilopoulou, S., & Talving, L. (2024). Euroscepticism as a syndrome of stagnation? Regional inequality and trust in the EU. *Journal of European Public Policy*, 31(6), 1494–1515. <https://doi.org/10.1080/13501763.2023.2264891>
- Vasilopoulou, S., & Zur, R. (2024). Electoral competition, the EU issue and far-right success in Western Europe. *Political Behavior*, 46(1), 565–585. <https://doi.org/10.1007/s11109-022-09841-y>
- Wassenberg, B. (2010). Can cross-border corporation provoke opposition to Europe? The example of the Upper Rhine Region. In B. Wassenberg, F. Clavert & P. Hamman (Eds.), *Contre L'Europe? Anti-européisme, euroscepticisme et alter-européisme dans la construction européenne de 1945 à nos jours* (Vol. 1, pp. 285–305). Steiner.
- Wassenberg, B., & Reitel, B. (Eds.). (2020). *Critical dictionary on borders, cross-border cooperation and European integration*. Peter Lang. <https://doi.org/10.3726/b15774>
- Weber, F. (2022). Cross-border cooperation in the border region of Germany, France, and Luxembourg in times of Covid-19. *European Societies*, 24(3), 354–381. <https://doi.org/10.1080/14616696.2022.2076894>
- Wheatley, J. (2016). Cleavage structures and dimensions of ideology in English politics: Evidence from voting advice application data. *Policy & Internet*, 8(4), 457–477. <https://doi.org/10.1002/poi3.129>

About the Authors



Daniela Braun is a professor of political science at Saarland University (USAAR). Her research interests include the politics of the European Union, party politics, public opinion, and political behaviour. She is currently involved in two Horizon Europe projects: ActEU and UNTWIST; she is also the co-lead on the SOUNDS project.



Elisa Deiss-Helbig is a research fellow at the Zukunftskolleg and a researcher in the Department of Politics and Public Administration at the University of Konstanz. Her research explores how electoral promises and their implementation contribute to inequalities in the representation of social groups, and how voters respond to group-targeted electoral appeals. Ph: Ines Janas



Theresa Gessler is a professor of political science at the University of Hamburg, with a focus on democratic political decision-making. Her research focuses on democracy, digital politics, gender, and immigration, often through the lens of party politics.



Jochen Müller is a professor of political sociology and methods at the University of Greifswald. His research focuses on legislative behaviour, party competition, and political attitudes and behaviour. His work has been published in journals such as *Electoral Studies*, *Legislative Studies Quarterly*, *Political Analysis*, and *Regional Studies*.



Julia Wagner is a doctoral candidate at the Chair for Comparative European Politics at Saarland University. She is working on a BMBF project focusing on the strategic communication of uncertainty in simulation models (STRAKOSIM). Ph: RPTU Thomas Koziel



Georg Wenzelburger is a professor of comparative European politics at Saarland University. He studies comparative public policy in Western Europe, focusing on the politics of insecurity, welfare state reform, and border regions. Recent publications appeared in the *European Journal of Political Research*, *West European Politics*, and *Regional and Federal Studies*.

Measuring Issue Congruence in Western Europe: How Voting Advice Applications Compare to Expert- and Manifesto-Based Estimates

Mattia Gatti 

Department of Political Science, LUISS University, Italy

Correspondence: Mattia Gatti (mgatti@luiss.it)

Submitted: 31 August 2025 **Accepted:** 21 January 2026 **Published:** 4 March 2026

Issue: This article is part of the issue “Voting Advice Applications: Methodological Innovations, Behavioural Effects, and Research Perspectives” edited by Diego Garzia (University of Lausanne / Bologna University), Stefan Marschall (Heinrich Heine University Düsseldorf), Mathias Wessel Tromborg (Aarhus University), and Andreas Albertsen (Aarhus University), fully open access at <https://doi.org/10.17645/pag.i485>

Abstract

Do congruence estimates derived from different sources converge? A growing number of scholars are studying party–voter congruence in a comparative perspective, increasingly focusing on specific issues that depart from the general left–right dimension. When doing so, however, they often encounter a missing-data problem: Researchers must combine citizen preferences and party positions derived from distinct sources, frequently presenting different question formats, wordings, and scales. This exacerbates perceptual biases and requires demanding assumptions regarding the common understanding of issues between voters and experts. While multiple techniques have been proposed to address these issues, they remain burdensome and rarely used. Consequently, this article focuses on the most widely employed approaches to estimating party–voter issue congruence. I confront measures constructed by matching European Election Study voter preferences with expert (Chapel Hill Expert Survey) and Euromanifesto party position estimates against those generated by voting advice applications (VAAs)—specifically the euandi dataset—which present the benefit of measuring parties and voters on the same scale, with the same wording. Focusing on 60 party–voter dyads across 12 Western European countries in 2014, the results indicate clear divergence between estimates derived from manifesto data and those from VAAs. Correlations between VAAs and experts are consistently positive and significant, though moderate. These estimates do not seem to be driven by different sampling methods or quality filtering. Finally, an exploratory analysis on the sources of divergence finds that manifesto length, expert uncertainty, and party newness help explain some of the differences between congruence estimates.

Keywords

issue congruence; representation; voter–party linkages; voting advice applications

1. Introduction

Among the most influential normative models of representative democracy, the responsible party model (Schmitt & Thomassen, 1999) assumes that political representation takes place when the policy preferences of the citizenry are closely reflected in the policies enacted by their parties in office. To engender this linkage, parties are expected to offer clear and distinct policy options that align with citizens' electoral demands (Belchior, 2013). Equally important, voters should select the party that advocates policy programs that most closely align with their preferences (Thomassen & van Ham, 2014).

Congruence, defined here as the positional correspondence between voters and their preferred party, constitutes, thereby, an invaluable instrument to evaluate effective policy representation, as a close match between the representatives and the represented should signal a higher chance that the former will incorporate their voters' interests and demands into the decision-making process (Andeweg & Thomassen, 2005; Mansbridge, 2009).

Unsurprisingly, a rich tradition has flourished in the political representation literature examining the extent of this match, with a growing focus on the cross-national level (Belchior, 2013; Dalton, 2017b; Ezrow & Xezonakis, 2011). While the traditional approach examines ideological congruence on the overarching left-right dimension (e.g., Schmitt & Thomassen, 1999), the increasing salience of specific issues has prompted recent contributions to focus extensively on what Jacques Thomassen calls the "blind corner" of political representation research (Thomassen, 2012): the paucity of works measuring issue congruence. One of the main findings of this strand is that parties are considerably out of step with their own voters on policy issues that poorly relate to the ideological dimension (Elsässer & Schäfer, 2023). This is the case for immigration, European integration, gender issues, and the environment.

Despite its heightened relevance, the study of issue congruence presents significant methodological hurdles for researchers interested in comparing estimates across countries. This often forces researchers to combine voter issue preferences and party issue positions derived from distinct data sources, frequently measured at different points in time, on different scales, and presenting different question wordings (Rosset & Stecker, 2019). Although techniques exist that mitigate some of these issues (Aldrich & McKelvey, 1977; Kurella & Rapp, 2024; Park et al., 2006), such methods are often too data-intensive and cumbersome to apply. As a result, the issue congruence literature still predominantly relies on the practice of matching different data sources, notwithstanding the potential drawbacks.

As the issue congruence literature expands, along with the sources and methods to measure it, it thus becomes critical to understand how these heterogeneous estimates compare. In short, I ask: To what extent do congruence estimates derived from different sources converge? And, if divergence is found, what predictors help explain it?

To answer these questions, I first confront congruence estimates on five relevant issues by matching cross-national survey data on citizens' issue preferences with party positions derived from two sources—manifesto data and expert surveys—across 12 Western European countries in 2014. More specifically, I leverage data on voter issue preferences derived from the extensively used European Election Study (EES) 2014 Voter Study (Schmitt et al., 2015). I then separately match these data with party issue position

estimates from the 2014 Euromanifesto project (Schmitt et al., 2016) and the 2014 wave of the Chapel Hill Expert Survey (CHES; Polk et al., 2017). Innovatively, I assess the convergence (or better, divergence) of these more traditional measures with voting advice applications (VAAs)-generated data, specifically leveraging the 2014 euandi dataset (Garzia et al., 2017)—a pan-European VAA aimed at helping citizens make an informed choice in the 2014 European Parliament elections—containing party and voter preferences on the same issues, same scale, and at the same point in time. Going a step further, I explore the determinants of such divergence by investigating characteristics inherent to both the voter survey- and party-level.

Although this article cannot provide a test for convergent validation (Adcock & Collier, 2001) as no objective, drawbacks-free benchmark indicator exists, I argue that assessing the alignment of distinctly derived congruence estimates remains imperative. This is particularly so in light of the growing “representative deficit” characterizing Western European polities (Alvarez et al., 2014; Bright et al., 2020). Parties appear increasingly ill-suited to read and aggregate voter demands as a result of the decline of traditional collective constituencies and particularization of voter preferences—trends that have prompted a withdrawal from civil society toward collusive cartelization (Katz & Mair, 1995). In this context, extending the comparison to VAA-based congruence measures is particularly fruitful. Recent scholarship has demonstrated the applicability of VAAs for constructing congruence measures as well as their suitability for cross-national analysis (Costello et al., 2021). As contributions based on VAAs proliferate (Stockinger et al., 2024), this study advances the literature by exploring the extent to which measures derived from such sources converge with those constructed from more “traditional” methods.

I find manifesto-based issue congruence estimates to correlate moderately and inconsistently across the five issues. Conversely, I find higher levels of correlation as well as consistent significance between euandi and the CHES-based measures. The overlap appears higher on EU integration and redistribution, the latter being poorly convergent for both CHES and euandi with Euromanifesto. The exploratory analysis finds some evidence supporting systematic bias between euandi-based measures and both Euromanifesto and CHES data. On the voter survey side, sampling procedure and quality filtering do not improve the match between the measures. Conversely, the length of the manifesto, expert uncertainty, and the newness of the party in the euandi dataset appear to explain some of the divergence.

Matching multiple issue congruence estimates reveals important patterns of convergence and divergence. Euromanifesto-based congruence measures correlate only moderately across the five issues and do so inconsistently in terms of statistical significance. By contrast, congruence measures derived from euandi exhibit higher correlation coefficients and consistently significant associations with CHES-based estimates. Substantively, overlap across data sources is greatest on EU integration and redistribution. Yet redistribution also emerges as an issue of particularly weak convergence between Euromanifesto-based measures and both expert- and euandi-based indicators. Exploratory analyses further suggest that divergence across data sources is not random but partly systematic. Indeed, I find some indications of bias between euandi-based measures and both Euromanifesto and CHES data. More specifically, on the voter-survey side, neither sampling procedures nor quality filtering substantially improve correspondence between congruence measures. Instead, features of the estimation procedures underlying party position measures—manifesto length, expert uncertainty, and party newness in the euandi dataset—account for at least some of the observed discrepancies. Taken together, these findings speak directly to both the congruence literature and the growing body of work leveraging VAAs to study political representation.

2. Combining Voter Preferences and Party Position Estimates Derived From Different Data Sources

Comparative studies on issue congruence often struggle with a missing-data problem. The traditional relevance of the left–right dimension as a super-issue structuring party–voter relations across Western Europe has not only contributed to confining studies on party–voter congruence to the realm of ideological dimensions (Belchior, 2013); it has markedly shaped the set of political attitudes and preferences that are measured through cross-national surveys.

In a context where multidimensional congruence—i.e., the ability of parties and governments to be in line with citizen preferences on different issues—has become an increasingly relevant indicator of the quality of democracy (Stecker & Tausendpfund, 2016), scholars wanting to estimate issue congruence in a comparative perspective must confront with two sets of methodological challenges: (a) finding survey items through which to gauge voter preferences on specific issues; (b) meaningfully combining these items with comparable estimates of party issue positions, ultimately generating a dataset that builds on two distinct sources.

2.1. Mass Survey Items Measuring Voters' Issue Preferences

The first step in measuring issue congruence between parties and voters is to retrieve information about their preferences on issue sets of interest. An examination of the empirical literature indicates that the large-scale cross-national European Social Survey (ESS; e.g., Devine & Ibenskas, 2021; Ibenskas & Polk, 2022; Rosset & Stecker, 2019; Stecker & Tausendpfund, 2016) and the EES (Belchior, 2013; Costello et al., 2014; Dalton, 2017a, 2021; Lefkofridi & Casado-Asensio, 2013; Rosset & Stecker, 2019) constitutes one of the most frequently used sources of voters' preferences in issue congruence studies focusing on Western European polities. While the EES is particularly well-suited, the ESS has also been widely employed to measure positional correspondence as it provides authoritative cross-national mass surveys on social and political values and attitudes since 2002, providing researchers with a longitudinal perspective on multi-dimensional congruence (see e.g., Devine & Ibenskas, 2021; Ibenskas & Polk, 2022; Rosset & Stecker, 2019; Stecker & Tausendpfund, 2016).

The EES Voter Study is a nationally representative post-election survey conducted in each member country of the EU in an arrangement with the European Parliament. Launched in 1979, it offers post-election individual-level survey data on electoral behaviour in both European Parliament and national elections, trust in and assessment of European institutions, as well as information on general value orientations. It is important to note that the EES Voter Study includes questions assessing voter preferences on various policy issues. These questions cover attitudes towards European unification and enlargement, state regulation and market control, wealth redistribution, taxation, public spending, same-sex marriage, civil liberties, immigration, and the environment. Additionally, since 2014, the EES Voter Survey has presented these questions using the same scale and wording as the CHES, one of Europe's most prominent sources of party position estimates.

2.2. Party Position Estimates

Plenty of cross-national party position estimates from different sources exist, yet they vary in the extent to which they can be suitably employed in measuring party–voter issue congruence. The two most widely used sources are manifesto data and expert-based judgments.

The former rely on hand-coding quasi-sentences which are assigned to—often—pre-defined categories of left and right. The position estimate is then operationalized as the proportion of left versus right quasi-sentences in the party manifesto. The reliance on party manifestos makes large-scale cross-national datasets such as the Manifesto Project (Lehmann et al., 2025) or the Euromanifesto project (Schmitt et al., 2016) extremely valuable for time series analyses, as some data can reach back to the pre-war period. Moreover, manifesto-based measures draw on written and publicly available records, allowing for competing and replicable measurement of party positions (Hooghe et al., 2010).

That said, combining manifesto data with voter preference data poses several challenges. For example, Marks et al. (2007) note how some manifestos may be too short to capture variation among parties and to provide precise measurements on issues that depart from the left–right dimension. This helps governing parties—which also tend to be older and gain more votes—be measured more accurately as they publish longer manifestos explaining their positions on a wide variety of issues (Gabel & Huber, 2000). Most importantly, manifesto data do not lend themselves too well to being combined with voter preference data (Kurella & Rapp, 2024). On the one hand, while, e.g., the ESS and EES present respondents with specific labels for the endpoints of a position item, the endpoints of the scale in, e.g., the Manifesto Project are just the maximum imbalance between two categories. On the other hand, the nature of the coding scheme makes it problematic to obtain fine-grained data on substantive issues. For example, the Manifesto Project provides two non-specific categories on immigration: Multiculturalism: Positive (607) and Multiculturalism: Negative (608). Similarly, quasi-sentences on European integration are clustered in just one category: Multiculturalism: Negative (608).

Against this background, it is unsurprising that—despite their continued use (e.g., Werner, 2020)—researchers seeking to combine party positions with voter issue preferences have increasingly turned to expert survey data. Here, party position estimates are obtained by averaging all expert judgments. Expert-based measures address several of the shortcomings of manifesto data by directly eliciting party positions on clearly defined policy scales, which are more readily comparable to voter survey items. In this approach, party positions are typically derived by aggregating expert judgments, most commonly through simple averaging, yielding estimates that are both substantively interpretable and cross-nationally comparable. Expert surveys have consistently produced reliable and valid measures of party positions when cross-validated with other measures (Hooghe et al., 2010; Marks et al., 2007). Importantly, what makes them suited for measuring issue congruence is their traditional focus on concrete issue dimensions. One example is provided by the Ray dataset (Ray, 1999), measuring party positions on EU integration and policies since 1984. This dataset has then been extended by the CHES, which to this day represents the most complete expert data source on party issue positions. In the 2019 wave of the CHES (Jolly et al., 2022), experts provide party positions on specific issues such as redistribution, state intervention in the economy, European integration, social lifestyle, immigration, and environment, as well as on more general issue dimensions as in the case of the GAL–TAN dimension (Hooghe & Marks, 2018). Further, experts in CHES

and other expert surveys are asked to evaluate positioning directly on structured scales with meaningful endpoints, allowing for a less problematic match with voter survey data.

2.3. Challenges in Matching Voter Issue Preferences With Party Position Estimates

As noted above, a large number of contributions in the issue congruence literature are nowadays matching voter preferences with party position estimates from either manifesto or expert data. Even if the selected items to be matched are comparable (e.g., same wordings and scales), researchers face a serious methodological challenge whose severity resides in perceptual biases and information asymmetry between experts and the average voter. As long noted by the literature (Aldrich & McKelvey, 1977), voter responses to issue scales and their perceived party positions suffer from differential item functioning (DIF). DIF arises when survey respondents perceive survey items in fundamentally different ways. Golder and Stramski (2010) highlight how this perceptual bias can be markedly accentuated in comparative analysis settings. That is because respondents in different countries may place themselves differently on a scale even though they share identical preferences, or they may place themselves at the same point on the scale even though the nature of political debate in these countries is different (King et al., 2004). The latter often describes the fact that respondents tend to locate themselves—or their preferred party—toward the centre of the scale, even when they—or their preferred party—are non-centrist, while pushing disliked parties toward the extremes. This is because moderation is generally considered a socially desirable trait, implying being reasonable and objective (Hare et al., 2015). The main consequence of this pattern of social desirability is an underestimation of polarization in voter opinions.

While voters tend to prime moderation over extremism, experts might comfortably locate parties clearly at the endpoints of the scale. While there is some empirical evidence challenging the assumption of expert judgments being free from perceptual biases (Curini, 2010), experts do tend to interpret issue scales differently from voters. They are usually more politically interested than the average voter and have access to a variety of different sources such as opinions voiced by intra-party factions (Marks et al., 2007). In turn, they are likely to interpret the endpoints of the scale more consistently through dimensions and time, making their judgments less amenable to those short-term factors (e.g., current political debates) affecting survey respondents (Kurella & Rapp, 2024, p. 3).

Several methods have been proposed to eliminate these systematic sources of bias, all sharing Aldrich and McKelvey's (1977) intuition to estimate respondent's distortion parameters (the intercept/"shift" and weight/"stretch") to recover the locations of voters and parties on a common underlying dimension (Hare et al., 2015; Jessee, 2021; Kurella & Rapp, 2024). Nonetheless, besides the rescaling technique developed by Kurella and Rapp (2024), these methods require voters' perceived party positions, which are seldom asked on concrete issues, especially from a comparative perspective.

2.4. VAAs

VAAs represent an emerging method for studying political representation in Europe. First introduced in the late 1990s, these informational devices help voters find the "best party fit" based on the match between their recorded policy issue preferences and the issue positions held by a set of political parties. Importantly, their two-pronged nature allows researchers to estimate issue congruence measures without resorting to two distinct data sources (Garzia et al., 2017).

Consequently, a primary advantage of VAAs is the measurement of voter and party preferences on the same issues, using the same scale and wording, and at the same point in time (Garzia & Marschall, 2019). This allows scholars to rely on fewer demanding assumptions regarding the shared understanding of items between experts and voters, as both groups are presented with identical scale endpoints and labels. A further benefit stems from the specific methodology frequently employed by these applications to estimate party positions. The iterative (Kieskompas) method combines the insights from country teams of experts and an extensive textual analysis with parties' self-placements on each policy item included in the survey (Ferreira da Silva et al., 2023). Including parties in the positioning process should minimize some of the weaknesses typical of methods based solely on expert judgments, namely informational asymmetry regarding different political parties. As noted by Marks et al. (2007), errors in expert surveys such as CHES are most pronounced for new, small, and non-governmental parties. These parties tend to garner fewer votes, be less visible in the media, and switch positions more readily, making it harder for experts to locate these parties accurately (Steenbergen & Marks, 2007). Finally, by utilizing parties' self-placements, VAAs reduce situations in which party position estimates are derived from only a small number of expert judgments, whose scores may additionally be in sharp disagreement (Hooghe et al., 2010; Marks et al., 2007).

Concurrently, VAA user data present distinct advantages over traditional surveys of public opinion. Because they attract an ever-increasing number of respondents during election campaigns, VAAs provide researchers with sufficient observations to reliably estimate the aggregate preferences of smaller parties' electorate base, usually plagued by a small-*n* problem (McDonnell & Werner, 2019).

To be sure, VAA-generated data are not without limitations. First, regarding measurement validity, social desirability bias remains a challenge even with identical scales. Additionally, research indicates that party self-positioning in VAAs may respond to strategic incentives, introducing noise into the equation (Cedroni & Garzia, 2010; Ilmarinen et al., 2022). Further, being freely available online, these tools suffer from self-selection bias (Pianzola, 2014). This usually results in convenience samples featuring younger and better-educated respondents who are also highly interested in politics (Garzia & Marschall, 2014). As noted by Bright et al. (2020), the highly non-representative nature of the sample can be mitigated by recalling Almond's (1950) seminal distinction between the general public (the object of traditional representative samples in election and social surveys) and the attentive public largely informing the former through a process of osmosis. Even so, this does not eliminate the non-representativeness of the sample, which must necessarily be addressed to produce reliable congruence estimates.

Notwithstanding these limitations, VAA data significantly enhance our understanding of public opinion while facilitating a more nuanced study of political representation (De Angelis, 2020; Popp, 2024). Recently, scholars have begun utilizing VAAs to derive congruence measures (Costello et al., 2021), a development paralleled by the shift from purely national applications to cross-national platforms. A prime example is the euandi dataset (Garzia et al., 2015). Developed by the European University Institute of Florence, it represents the first cross-national VAA project cumulating users' policy preferences and parties' positions across three European Parliament elections (2009, 2014, 2019). The dataset covers 27 EU countries and the UK, including all parties credited with at least 1% of the popular vote in the opinion polls preceding the respective European Parliament election (Ferreira da Silva et al., 2023). Both parties and users are provided with a set of identical policy statements. These five-point Likert scales measure the level of (dis)agreement with the respective statements: E.g., the item measuring users' and party positions on welfare benefits says,

“Social programmes should be maintained even at the cost of higher taxes,” where 1 means *completely disagree* and 5 *completely agree*.

In sum, VAAs present an emerging alternative to traditional methods of studying political representation across Europe. An ever-growing number of contributions now match users’ and parties’ information provided by VAAs to examine the “representative deficit” in Western democracies (e.g., Bright et al., 2020; McDonnell & Werner, 2019; Popp, 2024). Further, growing evidence attests to the reliability and validity of VAAs’ party position estimates vis-à-vis other standard methods (Ferreira da Silva et al., 2023; Gemenis, 2013).

What remains absent, however, is a systematic comparison of VAA-generated party–voter issue congruence measures with more established approaches to measuring congruence—specifically, those combining election surveys with expert or manifesto data. The remainder of this article is dedicated to addressing this gap.

3. Data and Methods

The first step in analysing congruence is choosing how to conceptualize it and measure it. As noted by Golder and Stramski (2010, p. 92), I can think of congruence as concerning one citizen and one representative (one-to-one relationship), many citizens and one representative (many-to-one relationship), and many citizens and many representatives (many-to-many relationship). In this article, I define congruence as a many-to-one relationship involving parties and their voters, thereby incorporating information about the distribution of citizen preferences. Accordingly, issue congruence is measured as the average distance between a party and its voters, so that each respondent’s congruence score is averaged and then aggregated per party. Crucially, I depart from previous contributions measuring “absolute” issue congruence (Giger & Lefkofridi, 2014). Instead, I follow Rosset and Stecker (2019), and I operationalize congruence in relative terms. This allows me to assess the direction of (in)congruence between parties and voters, providing a more refined picture of the convergence—or rather divergence—of congruence estimates.

The formula I use is the following:

$$RCC = \frac{1}{N} \sum_{i=1}^N P_i - C_i$$

Where N is the number of citizens having supported party P_i , and C_i is the ideal point of citizen i^{th} .

I combine four distinct data sources—the 2014 EES Voter Study, the 2014 Euromanifesto project, the 2014 CHES, and the 2014 euandi dataset—to generate three sets of issue congruence estimates: (a) EES-Euromanifesto, (b) EES-CHES, and (c) euandi. Table A1 in the Supplementary File provides the list of issue position items and question wording. To allow for meaningful matching, I recoded all items into the same polarity so that higher values represent a more market-liberal economic preference/position or a more conservative stance on cultural issues. Further, I normalized all response categories on a 0–1 scale. Accordingly, negative issue congruence scores signify that parties hold a relatively more left-wing position than their voters. Conversely, positive values highlight a relatively more right-wing party position than their electorate.

I restrict my analysis to five issues: redistribution, immigration, EU integration, gender issues, and the environment. While euandi offers a comprehensive range of policy statements, the EES and CHES are more limited in their coverage, necessitating this focus. Nevertheless, these five issues provide a comprehensive mapping of the Western European political space, capturing the increasing relevance of secondary issue dimensions (Hooghe & Marks, 2018; Kriesi et al., 2006).

Except for the euandi dataset, I used binary vote choice items to match parties and the respondents who stated they have voted for them either at the most recent national election or at the latest European Parliament election. For the former, I leveraged propensity-to-vote (PTV) scores, reported on a 10-point scale. Where the user indicated the same PTV value for more than one party, I selected the party with the highest agreement score. To ensure comparability with the VAA estimates, which are inherently tied to the 2014 electoral context, I utilize the EES item on European Parliament vote choice. For reliability purposes, I excluded parties with less than 30 supporters for either of the combined data sources. The final dataset returns 60 party–voter dyads. My analysis extends to 12 Western European countries: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, and Sweden.

Several important methodological issues arise when comparing estimates derived from such heterogeneous sources and must be considered when interpreting the findings. Whereas the composite congruence measures rely on retrospective vote recall (EES), the euandi data utilize prospective vote intentions. Although this can induce divergence due to cognitive biases (van Elsas et al., 2016), the immediate fieldwork of the EES following the 2014 European election partially mitigates this concern (Schmitt et al., 2015). Second—similarly—the timing of data collection introduces potential noise. Notably, CHES data were collected well after the election, with experts evaluating party positions for the entire year rather than the specific campaign period. Third, due to the different wording and scale endpoints of the items on which party and voters are matched, there is no certainty that we are not comparing “apples” with “pears.” This problem is exacerbated by the fact that euandi’s issue statements are highly detailed and positional in nature, compared to the higher level of abstraction on which CHES operates. While it is more difficult to make assumptions about voters, Kurella and Rapp (2024) show how party standpoints on different issues from the same underlying dimension (economic or cultural) are highly correlated. I assume this to be the case between party–issue items of CHES and euandi belonging to the same underlying dimension.

4. Empirical Results

4.1. *Confronting Three Sets of Congruence Estimates*

Table 1 reports pairwise correlation coefficients between congruence estimates derived from the three data sources across five issues. Focusing first on issue-specific patterns, only two issues—European integration and immigration—display consistently positive and statistically significant correlations across all three datasets. This pattern suggests a certain degree of convergence across both sources and methods. By contrast, congruence estimates for the remaining issues show more inconsistency, both in terms of magnitude and statistical significance across datasets. Redistribution is a particularly salient case: Correlations are weak and, in two out of three pairwise comparisons, negative and statistically insignificant.

Such inconsistencies may point to the fact that different estimation strategies ultimately capture distinct aspects of issue congruence rather than a single underlying construct. This interpretation is reinforced when examining the specific matching between estimates. The convergence between manifesto-based congruence measures and euandi estimates is limited: Beyond European integration and immigration, no additional issue reaches conventional levels of statistical significance. A similar, though somewhat less pronounced, pattern emerges when comparing Euromanifesto-based measures with those derived from CHES. Apart from European integration, correlations across the remaining issues are at best moderate.

Importantly, I observe consistently moderate-to-strong convergence between congruence estimates derived from EES-CHES and those based on euandi. Both the stability and magnitude of these correlations suggest these estimates likely measure similar things. This greater convergence is plausibly attributable to the substantial similarities between the two methods. In both cases, party positions are ultimately grounded in expert-based judgements: euandi's iterative method can be understood as a more refined extension of traditional expert surveys, combining expert judgments with parties' self-placement. Importantly, as shown in Table A10 in the Supplementary File, the correlation coefficients remain equally strong and statistically significant when I substitute EES voter data with estimates derived from the ESS.

By contrast, the strong overlap between EES-CHES and euandi, coupled with the markedly weaker correlations involving Euromanifesto-based congruence estimates, lends indirect support to earlier work identifying manifesto data as something of an outlier among party positioning methods (Kurella & Rapp, 2024; Rohrschneider & Whitefield, 2012). A key reason lies in the fundamentally different logic underpinning manifesto-based measures. Whereas CHES and euandi ask experts to locate parties directly on specific issue scales, projects such as the Euromanifesto or the Manifesto Project infer party positions from the relative salience of issue categories within manifestos. As a result, estimated positions on any given issue or broader dimension are affected by the salience of other categories in the document, reducing the likelihood of extreme placements and, in turn, convergence with expert-based measures (Ferreira da Silva et al., 2023, p. 17). This remains the case even when, as in my analysis, I apply the logit-scaling technique proposed by Lowe et al. (2011), which addresses several well-known limitations often associated with manifesto-based data.

Table 1. Correlations between congruence measures constructed from alternative datasets by issue.

Issue	Euromanifesto/CHES	N	Euromanifesto/euandi	N	CHES/euandi	N
Redistribution	−0.11	60	−0.09	59	0.56***	59
Immigration	0.36***	60	0.22*	58	0.44***	58
EU integration	0.68***	60	0.63***	60	0.72***	60
Gender issues	0.41***	60	0.01	60	0.46***	60
Environment	0.25*	60	0.17	57	0.27**	57

Notes: Pearson correlations (r); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Ns are pairwise complete observations.

Figure 1 illustrates these patterns by presenting scatterplots of the congruence estimates. When pooling all five issues, the strongest correlation is observed between the EES-CHES and euandi measures ($r = 0.53$). This is closely followed by the correlation between the two EES-based estimates (EES-CHES and EES-Euromanifesto), which stands at $r = 0.47$. By contrast, the convergence between euandi and Euromanifesto-based congruence is markedly lower ($r = 0.26$), further highlighting the distinctiveness of the manifesto-derived data.

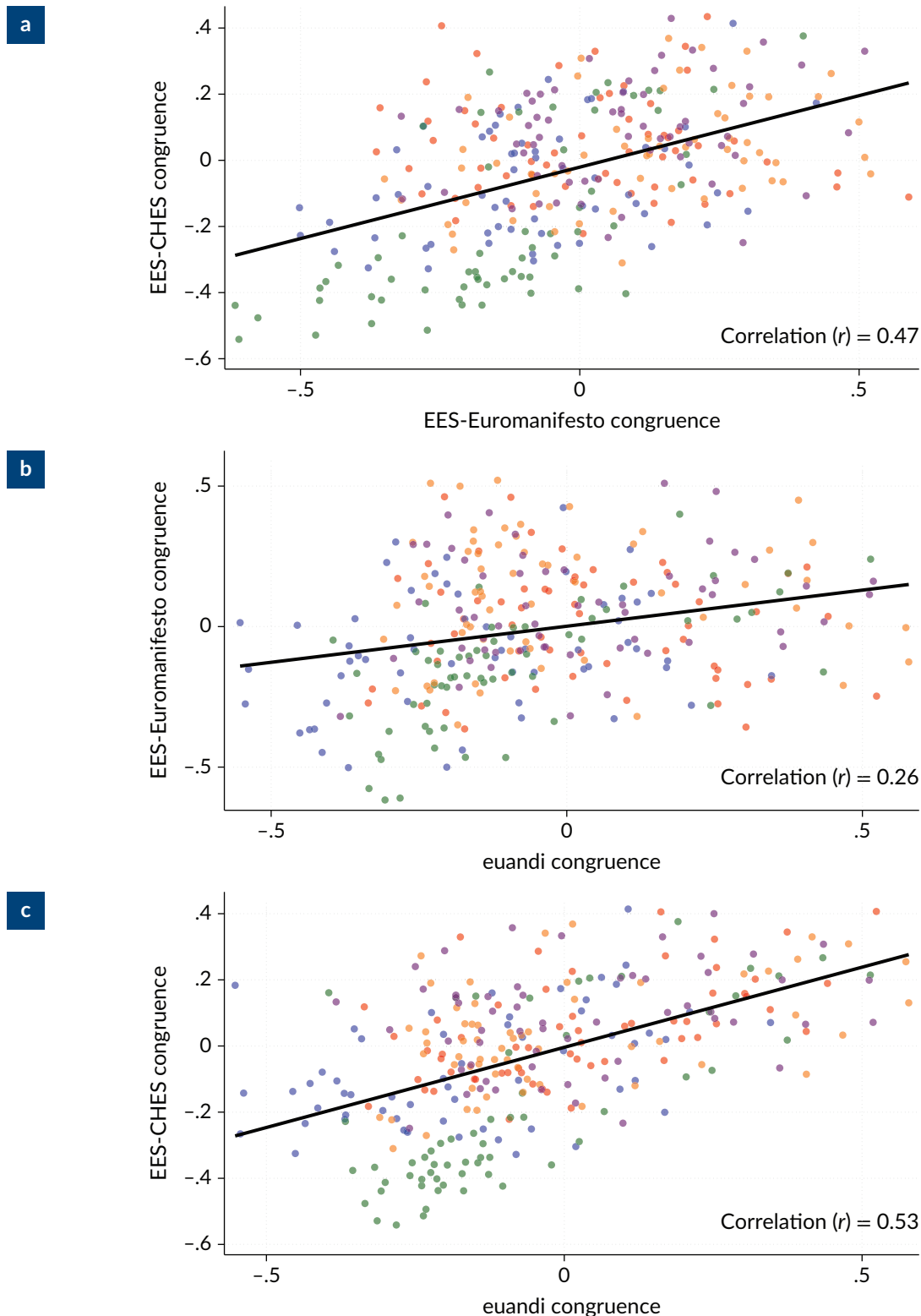


Figure 1. Overall correspondence between alternative issue congruence measures: (a) EES-Euromanifesto and EES-CHES congruence estimates; (b) euandi and EES-Euromanifesto congruence estimates; (c) euandi and EES-CHES congruence estimates. Notes: The panels report pairwise correlations between congruence estimates derived from three data combinations, EES-Euromanifesto, EES-CHES, and euandi-based measures; each panel plots one congruence estimate on the horizontal axis against another on the vertical axis; the pooled sample comprises 300 observations across 12 countries and five issues.

4.2. Exploring Divergence: Sampling Techniques and Quality Filtering

Moving a step further, I ask which specific features of the composite methods used to derive the congruence measures help explain at least part of the observed differences across data sources. A natural starting point is the respondent side of the equation, and particularly the sampling methodologies underlying the different datasets.

As discussed above, whereas the EES relies on probability sampling, euandi draws on a self-selected subset of voters who are disproportionately male, younger, more highly educated, and politically sophisticated (van de Pol et al., 2014). If self-selection affects the distribution of voter preferences, it may in turn translate into differences in estimated levels of issue congruence. To assess whether such selection effects contribute to the divergence between congruence measures—and to test whether correcting for sample composition improves the performance of VAA-based estimates—I follow Toshkov and Romeijn (2021) and apply post-stratification weights to the euandi data. Specifically, observations from underrepresented demographic groups are up-weighted, while those from overrepresented groups are down-weighted. The weighting scheme is calibrated using the demographic profiles of party electorates derived from the same 2014 EES. Tables A2 to A4 in the Supplementary File report the demographic profile (gender, age, and education level) of the EES, euandi, and ESS.

Table 2 reports the pairwise correlations between the weighted euandi congruence estimates and those derived from EES-Euromanifesto and EES-CHES data. Contrary to expectations, weighting does not lead to a meaningful improvement in correspondence across measures: Convergence with Euromanifesto-based estimates remains weak, while correlations with expert-based measures continue to be comparatively strong.

Table 2. Correlations between congruence measures constructed from alternative sources by issue with weighted euandi user data.

Issue	Euromanifesto/euandi	N	CHES/euandi	N
Redistribution	−0.07	59	0.56***	59
Immigration	0.27**	58	0.45***	58
EU integration	0.64***	60	0.70***	60
Gender issues	0.11	60	0.41***	60
Environment	0.18	57	0.26*	57

Notes: Pearson correlations (r); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; N s are pairwise complete observations; euandi user data are weighted by age, gender, and education based on socio-demographic information derived from the EES Voter Study sample.

Beyond sampling, differences in the quality of respondent-level data may also contribute to variation in convergence across congruence estimates. First, previous research has shown that VAA top-PTV scores do not always correspond to the party ultimately chosen at the ballot box (Talukder et al., 2021; Walgrave et al., 2008). This evidence raises concerns about relying on agreement scores as a decisive criterion for retaining or excluding VAA-user observations, as such a strategy may introduce additional noise into congruence estimates. To assess the sensitivity of my results, I thus re-estimate correlations after excluding all observations with tied top-PTV scores (i.e., cases in which two or more parties share the highest score). Second, I exclude respondents who completed a VAA questionnaire for a country other than their country of residence, on the assumption that these individuals may be less familiar with the relevant party system.

Third, I remove observations associated with unusually short or long completion times, which may signal suboptimal engagement with the survey instrument. I operationalize this by excluding observations with completion times below the 10th percentile and above the 90th percentile.

Tables A5 to A8 in the Supplementary File report the pairwise correlations based on these adjusted samples. Across all specifications, the resulting coefficients closely resemble those reported in Table 2. Taken together, these analyses provide little evidence that sampling methodology or data-quality filtering on the respondent side substantively affects the level of convergence between euandi-based congruence measures and those derived from Euromanifesto- or expert-based data.

4.3. Exploring Divergence: Features of Party Position Estimation Methods

As a final step, I move to the party-level of the equation and explore potential determinants of divergence between congruence estimates that may ultimately derive from the way in which party position estimates are constructed or derived. To be sure, predicting divergence is not straightforward when examining composite measures such as issue congruence ones. With this in mind, I took inspiration from previous studies examining patterns of convergence across datasets (e.g., Ferreira da Silva et al., 2023) and operationalized divergence as the absolute difference in issue congruence for each party between pairs of datasets: EES-Euromanifesto and EES-CHES, euandi and EES-Euromanifesto, and euandi and EES-CHES, calculated separately for each of the five issues. This procedure yields a continuous measure capturing the overall degree of disagreement between methods, with higher values indicating greater divergence. Using these measures as dependent variables, I then estimated three separate OLS regressions, one for each combination of data sources.

I include several predictors corresponding to the specific estimation procedures of each congruence measure. For Euromanifesto-based measures, I use a continuous variable counting the number of quasi-sentences devoted to the analysed issues, as shorter manifestos with fewer statements are likely to yield less precise estimates. For expert-based measures, I include a measure of expert uncertainty, calculated as the sum of the standard deviations of expert judgments across the five issues. Higher disagreement among experts reduces confidence in the mean position assigned to a party. For VAA-based measures, I include a measure of party newness, indicating whether the party has never been included in previous waves of the euandi dataset, and a dummy variable identifying whether the party actively participated in the iterative process, placing itself on the issues. In addition, I consider two characteristics that may affect all methods equally: Party size and government status are expected to improve the quality of party position estimates, as larger and governing parties are more visible and better documented and tend to have a more extensive and diverse policy agenda (Ferreira da Silva et al., 2023; Marks et al., 2007).

Figure 2 presents the coefficient plot from the OLS regression analysis explaining the divergence between the two “traditional” measures: manifesto- and expert-based congruence estimates. As the figure illustrates, no predictor appears to systematically affect the alignment between these two sources.

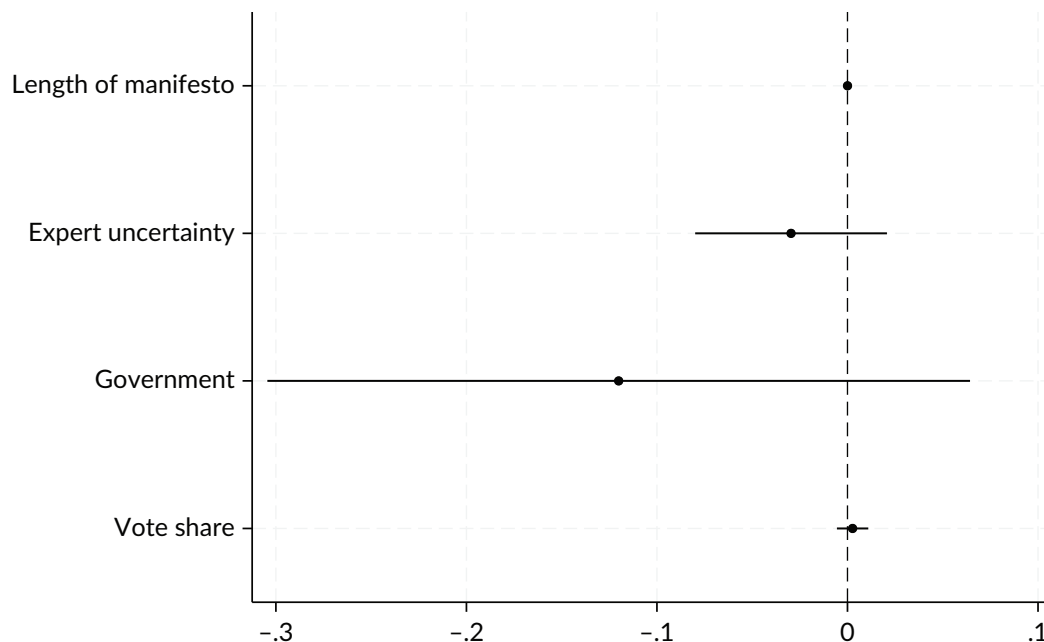


Figure 2. Determinants of divergence between EES-Euromanifesto and EES-CHES congruence estimates (coefficient plot). Notes: The figure displays OLS regression coefficients with 90% confidence intervals; the dependent variable is the absolute difference between congruence estimates derived from EES-Euromanifesto and EES-CHES; horizontal lines crossing the vertical zero line indicate statistically insignificant effects.

Turning to Table 3, however, I find evidence that the divergence involving euandi-based estimates is at least partially systematic. Model 1, which analyses the discrepancy between euandi and Euromanifesto, reveals a negative and significant relationship between manifesto length and divergence. Specifically, the higher the number of quasi-sentences in a party manifesto, the lower the difference (i.e., the higher the convergence) between the two measures. This suggests that longer manifestos—which likely provide more granular policy information—improve the accuracy of position estimates.

In contrast, Model 2 identifies expert uncertainty as a driver of divergence between CHES and euandi measures. I find that higher levels of standard deviation among experts regarding a party’s position significantly led to larger discrepancies ($b = 0.05$, $p < 0.05$). This highlights a potential limitation of expert-based approaches: When experts struggle to agree on a placement—likely also due to the complexity of comparing parties across countries or time—the resulting estimates tend to drift further from the VAA benchmark.

Finally, regarding shared factors, being a “new party” in the euandi dataset leads to significantly greater divergence ($b = 0.55$, $p < 0.05$), but this effect is confined to the Euromanifesto comparison (Model 1). This likely reflects the difficulty of coding shorter or less standardized platforms typical of emerging parties. Finally, I observe a weak, counterintuitive effect of incumbency: Being in government tends to slightly increase rather than decrease the divergence between estimates, specifically in the CHES–euandi comparison ($b = 0.15$, $p < 0.10$). Party self-placement, by contrast, shows no significant effect in reducing divergence across either model.

Table 3. OLS regression models explaining divergence between euandi and traditional congruence estimates.

	Diff. Euromanifesto/euandi (1)	Diff. CHES/euandi (2)
Length of manifesto	−0.00* (0.00)	
Expert uncertainty		0.05** (0.02)
New party	0.55** (0.27)	−0.06 (0.20)
Self-placement	0.02 (0.14)	0.04 (0.09)
Government	0.13 (0.14)	0.15* (0.09)
Vote share	0.00 (0.01)	0.00 (0.00)
R^2	0.14	0.20
Observations	57	55

Notes: Standard errors in parentheses; the dependent variable is the absolute difference between the congruence estimates derived from the specified sources; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5. Conclusion

Assessing the quality of political representation requires reliable measures of the linkage between citizens and their representatives. As the political space in Western Europe becomes increasingly multidimensional, with specific salient issues structuring party–voter relations, the need for precise issue congruence estimates is more acute than ever. However, scholars attempting to measure these linkages frequently encounter a missing-data problem, forcing them to combine heterogeneous data sources that often present different scales, wordings, and estimation techniques.

By systematically confronting “traditional” composite measures (EES-Euromanifesto and EES-CHES) with estimates generated by VAAs (specifically the euandi dataset), this study offers a critical comparison of these approaches. My analysis of 60 party–voter dyads across 12 countries yields three main insights.

First, methodological choices entail important consequences for the estimation of congruence. I find that Euromanifesto-based estimates correlate only moderately and inconsistently with other congruence measures. This reinforces the argument that the salience-based logic of manifesto coding translates imperfectly into the positional one required for party–voter comparisons. Conversely, the convergence between euandi and EES-CHES estimates suggests that these two methods measure common latent concepts. This likelihood is underpinned by their shared reliance on expert judgment.

Second, my exploratory analysis clarifies the sources of the observed divergence. Contrary to expectations regarding the “demand side,” the self-selected nature of VAA users does not appear to drive the discrepancies. Weighting the sample to correct for socio-demographic biases yielded no significant improvement in the match between congruence measures. Instead, characteristics inherent to the estimation of party positions—specifically manifesto length, expert uncertainty, and party newness—appear

to explain a portion of the divergence. This suggests that “noise” in congruence estimates often stems from the difficulty of locating parties (especially newer or smaller ones) rather than from the non-representativeness of the VAA user base.

These findings attest to the centrality of VAAs not merely as information tools for voters, but as a valuable data source for studying political representation. Having said that, euandi represents only one VAA, as the proliferation of the latter across Europe offers a rapidly expanding avenue for research and validation of these findings. Future scholarship should extend this validation to other platforms and election cycles, further refining scholars’ ability to monitor the “representative deficit” in an era of transforming party–voter linkages.

Acknowledgments

I would like to express my sincere gratitude to the three reviewers whose insights have been invaluable.

Funding

Publication of this article in open access was made possible through the institutional membership agreement between LUISS University and Cogitatio Press.

Conflict of Interests

The author declares no conflict of interests.

Data Availability

The data used are available at the following links:

https://search.gesis.org/research_data/ZA5917?doi=10.4232/1.12138

https://search.gesis.org/research_data/ZA5160

https://search.gesis.org/research_data/ZA5162

<https://ess.sikt.no/en/study/ccd56840-e949-4320-945a-927c49e1dc4f>

<https://www.chesdata.eu/2014-chapel-hill-expert-survey>

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529–546. <https://doi.org/10.1017/S0003055401003100>
- Aldrich, J. H., & McKelvey, R. D. (1977). A method of scaling with applications to the 1968 and 1972 presidential elections. *American Political Science Review*, 71(1), 111–130.
- Almond, G. A. (1950). *The American people and foreign policy*. Harcourt, Brace and Company.
- Alvarez, R. M., Levin, I., Mair, P., & Trechsel, A. (2014). Party preferences in the digital age: The impact of voting advice applications. *Party Politics*, 20(2), 227–236. <https://doi.org/10.1177/1354068813519960>
- Andeweg, R. B., & Thomassen, J. J. A. (2005). Modes of political representation: Toward a new typology. *Legislative Studies Quarterly*, 30(4), 507–528. <https://doi.org/10.3162/036298005X201653>

- Belchior, A. M. (2013). Explaining left–right party congruence across European party systems: A test of micro-, meso-, and macro-level models. *Comparative Political Studies*, 46(3), 352–386. <https://doi.org/10.1177/0010414012453695>
- Bright, J., Garzia, D., Lacey, J., & Trechsel, A. H. (2020). The representative deficit in different European party systems: An analysis of the elections to the European Parliament 2009–2014. *Italian Journal of Electoral Studies*, 83(1), 45–57. <https://doi.org/10.36253/qoe-9531>
- Cedroni, L., & Garzia, D. (2010). *Voting advice applications in Europe: The state of the art*. Civis.
- Costello, R., Thomassen, J., & Rosema, M. (2014). European Parliament elections and political representation: Policy congruence between voters and parties. In C. Arnold & M. Franklin (Eds.), *Assessing political representation in Europe* (pp. 10–32). Routledge.
- Costello, R., Toshkov, D., Bos, B., & Krouwel, A. (2021). Congruence between voters and parties: The role of party-level issue salience. *European Journal of Political Research*, 60(1), 92–113. <https://doi.org/10.1111/1475-6765.12388>
- Curini, L. (2010). Experts' political preferences and their impact on ideological bias: An unfolding analysis based on a Benoit-Laver expert survey. *Party Politics*, 16(3), 299–321. <https://doi.org/10.1177/1354068809341051>
- Dalton, R. J. (2017a). Citizens' representation in the 2009 European Parliament elections. *European Union Politics*, 18(2), 188–211. <https://doi.org/10.1177/1465116516689729>
- Dalton, R. J. (2017b). Party representation across multiple issue dimensions. *Party Politics*, 23(6), 609–622. <https://doi.org/10.1177/1354068815614515>
- Dalton, R. J. (2021). The representation gap and political sophistication: A contrarian perspective. *Comparative Political Studies*, 54(5), 889–917. <https://doi.org/10.1177/0010414020957673>
- De Angelis, A. (2020). How voters distort their perceptions and why this matters. In E. Suhay, B. Grofman, & A. H. Trechsel (Eds.), *The Oxford handbook of electoral persuasion* (pp. 945–973). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190860806.013.55>
- Devine, D., & Ibenskas, R. (2021). From convergence to congruence: European integration and citizen–elite congruence. *European Union Politics*, 22(4), 676–699. <https://doi.org/10.1177/14651165211024936>
- Elsässer, L., & Schäfer, A. (2023). Political inequality in rich democracies. *Annual Review of Political Science*, 26, 469–487. <https://doi.org/10.1146/annurev-polisci-052521-094617>
- Ezrow, L., & Xezonakis, G. (2011). Citizen satisfaction with democracy and parties' policy offerings. *Comparative Political Studies*, 44(9), 1152–1178. <https://doi.org/10.1177/0010414011405461>
- Ferreira da Silva, F., Reiljan, A., Cicchi, L., Trechsel, A. H., & Garzia, D. (2023). Three sides of the same coin? Comparing party positions in VAAs, expert surveys and manifesto data. *Journal of European Public Policy*, 30(1), 150–173. <https://doi.org/10.1080/13501763.2021.1981982>
- Gabel, M. J., & Huber, J. D. (2000). Putting parties in their place: Inferring party left-right ideological positions from party manifestos data. *American Journal of Political Science*, 44(1), 94–103. <https://doi.org/10.2307/2669295>
- Garzia, D., & Marschall, S. (2014). *Matching voters with parties and candidates: Voting advice applications in a comparative perspective*. ECPR Press. <https://cadmus.eui.eu/handle/1814/34156>
- Garzia, D., & Marschall, S. (2019). Voting advice applications. In W. R. Thompson (Ed.), *Oxford research encyclopedia of politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.620>
- Garzia, D., Trechsel, A., & De Sio, L. (2017). Party placement in supranational elections: An introduction to the euandi 2014 dataset. *Party Politics*, 23(4), 333–341. <https://doi.org/10.1177/1354068815593456>
- Garzia, D., Trechsel, A. H., De Sio, L., & De Angelis, A. (2015). *euandi: Project description and datasets documentation* (Research Paper No. RSCAS 2015/01). Robert Schuman Centre for Advanced Studies.

- 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>
- Giger, N., & Lefkofridi, Z. (2014). Salience-based congruence between parties & their voters: The Swiss case. *Swiss Political Science Review*, 20(2), 287–304. <https://doi.org/10.1111/spsr.12069>
- Golder, M., & Stramski, J. (2010). Ideological congruence and electoral institutions. *American Journal of Political Science*, 54(1), 90–106. <https://doi.org/10.1111/j.1540-5907.2009.00420.x>
- Hare, C., Armstrong, D. A., Bakker, R., Carroll, R., & Poole, K. T. (2015). Using Bayesian Aldrich-McKelvey scaling to study citizens' ideological preferences and perceptions. *American Journal of Political Science*, 59(3), 759–774. <https://doi.org/10.1111/ajps.12151>
- Hooghe, L., Bakker, R., Brigevid, A., De Vries, C., Edwards, E., Marks, G., Rovny, J., Steenbergen, M., & Vachudova, M. (2010). Reliability and validity of the 2002 and 2006 Chapel Hill expert surveys on party positioning. *European Journal of Political Research*, 49(5), 687–703. <https://doi.org/10.1111/j.1475-6765.2009.01912.x>
- Hooghe, L., & Marks, G. (2018). Cleavage theory meets Europe's crises: Lipset, Rokkan, and the transnational cleavage. *Journal of European Public Policy*, 25(1), 109–135. <https://doi.org/10.1080/13501763.2017.1310279>
- Ibenskas, R., & Polk, J. (2022). Congruence and party responsiveness in Western Europe in the 21st century. *West European Politics*, 45(2), 201–222. <https://doi.org/10.1080/01402382.2020.1859756>
- Ilmarinen, V.-J., Isotalo, V., Lönnqvist, J.-E., & von Schoultz, Å. (2022). Do politicians' answers to voting advice applications reflect their sincere beliefs? Comparing publicly and confidentially stated ideological positions in a candidate-centred electoral context. *Electoral Studies*, 79, Article 102504. <https://doi.org/10.1016/j.electstud.2022.102504>
- Jessee, S. A. (2021). Estimating individuals' political perceptions while adjusting for differential item functioning. *Political Analysis*, 29(1), 1–18. <https://doi.org/10.1017/pan.2019.47>
- Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M., & Vachudova, M. A. (2022). Chapel Hill Expert Survey trend file, 1999–2019. *Electoral Studies*, 75, Article 102420. <https://doi.org/10.1016/j.electstud.2021.102420>
- Katz, R. S., & Mair, P. (1995). Changing models of party organization and party democracy: The emergence of the cartel party. *Party Politics*, 1(1), 5–28. <https://doi.org/10.1177/1354068895001001001>
- King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191–207. <https://doi.org/10.1017/S000305540400108X>
- Kriesi, H., Grande, E., Lachat, R., Dolezal, M., Bornschie, S., & Frey, T. (2006). Globalization and the transformation of the national political space: Six European countries compared. *European Journal of Political Research*, 45(6), 921–956. <https://doi.org/10.1111/j.1475-6765.2006.00644.x>
- Kurella, A.-S., & Rapp, M. (2024). Combining voter preferences with party position estimates from different sources for studying voting behavior and representation. *Electoral Studies*, 87, Article 102734. <https://doi.org/10.1016/j.electstud.2023.102734>
- Lefkofridi, Z., & Casado-Asensio, J. (2013). European Vox Radicis: Representation and policy congruence on the extremes. *Comparative European Politics*, 11(1), 93–118. <https://doi.org/10.1057/cep.2012.1>
- Lehmann, P., Franzmann, S., Al-Gaddooa, D., Burst, T., Ivanusch, C., Regel, S., Riethmüller, F., Volkens, A., Weßels, B., & Zehnter, L. (2025). *The Manifesto data collection* (Version 2025a) [Data set]. Manifesto Project. <https://doi.org/10.25522/manifesto.mpps.2025a>
- Lowe, W., Benoit, K., Mikhaylov, S., & Laver, M. (2011). Scaling policy preferences from coded political texts. *Legislative Studies Quarterly*, 36(1), 123–155. <https://doi.org/10.1111/j.1939-9162.2010.00006.x>

- Mansbridge, J. (2009). A “selection model” of political representation. *Journal of Political Philosophy*, 17(4), 369–398. <https://doi.org/10.1111/j.1467-9760.2009.00337.x>
- Marks, G., Hooghe, L., Steenbergen, M. R., & Bakker, R. (2007). Crossvalidating data on party positioning on European integration. *Electoral Studies*, 26(1), 23–38. <https://doi.org/10.1016/j.electstud.2006.03.007>
- McDonnell, D., & Werner, A. (2019). Differently Eurosceptic: Radical right populist parties and their supporters. *Journal of European Public Policy*, 26(12), 1761–1778. <https://doi.org/10.1080/13501763.2018.1561743>
- Park, D. K., Gelman, A., & Bafumi, J. (2006). State-level opinions from national surveys: Poststratification using multilevel logistic regression. In J. E. Cohen (Ed.), *Public opinion in state politics* (pp. 209–228). Stanford University Press. <https://doi.org/10.11126/stanford/9780804753005.003.0011>
- 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>
- Polk, J., Rovny, J., Bakker, R., Edwards, E., Hooghe, L., Jolly, S., Koedam, J., Kostelka, F., Marks, G., Schumacher, G., Steenbergen, M., Vachudova, M., & Zilovic, M. (2017). Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Research & Politics*, 4(1). <https://doi.org/10.1177/2053168016686915>
- Popp, R. (2024). Harmony and dissonance: Unveiling issue linkages between voters and parties across EU democracies. *Party Politics*, 31(3), 476–487. <https://doi.org/10.1177/13540688231221650>
- Ray, L. (1999). Measuring party orientations towards European integration: Results from an expert survey. *European Journal of Political Research*, 36(2), 283–306. <https://doi.org/10.1111/1475-6765.00471>
- Rohrschneider, R., & Whitefield, S. (2012). *The strain of representation: How parties represent diverse voters in Western and Eastern Europe*. Oxford University Press.
- Rosset, J., & Stecker, C. (2019). How well are citizens represented by their governments? Issue congruence and inequality in Europe. *European Political Science Review*, 11(2), 145–160. <https://doi.org/10.1017/S1755773919000043>
- Schmitt, H., Braun, D., Popa, S. A., Mikhaylov, S., & Dwinger, F. (2016). *European Parliament Election Study 2014, Euromanifesto Study* (Version 1.0.0) [Data set]. GESIS Data Archive. <https://doi.org/10.4232/1.5162>
- Schmitt, H., Hobolt, S. B., Popa, S. A., & Teperoglou, E. (2015). *European Parliament Election Study 2014, Voter Study, First Post-Election Survey* (Version 1.0.0) [Data set]. GESIS Data Archive. <https://doi.org/10.4232/1.5160>
- Schmitt, H., & Thomassen, J. J. A. (Eds.). (1999). *Political representation and legitimacy in the European Union*. Oxford University Press.
- Stecker, C., & Tausendpfund, M. (2016). Multidimensional government-citizen congruence and satisfaction with democracy. *European Journal of Political Research*, 55(3), 492–511. <https://doi.org/10.1111/1475-6765.12147>
- Steenbergen, M. R., & Marks, G. (2007). Evaluating expert judgments. *European Journal of Political Research*, 46(3), 347–366. <https://doi.org/10.1111/j.1475-6765.2006.00694.x>
- Stockinger, E., Maas, J., Talvitie, C., & Dignum, V. (2024). Trustworthiness of voting advice applications in Europe. *Ethics and Information Technology*, 26(3), Article 55. <https://doi.org/10.1007/s10676-024-09790-6>
- Talukder, D., Uyttendaele, L., Jennart, I., & Rihoux, B. (2021). The impact of VAAs on vote switching at the 2019 Belgian legislative elections: More switchers, but making their own choices. *Politics of the Low Countries*, 3(1), 73–94. <https://doi.org/10.5553/PLC/.000010>
- Thomassen, J. (2012). The blind corner of political representation. *Representation*, 48(1), 13–27. <https://doi.org/10.1080/00344893.2012.653229>

- Thomassen, J., & van Ham, C. (2014). Failing political representation or a change in kind? Models of representation and empirical trends in Europe. *West European Politics*, 37(2), 400–419. <https://doi.org/10.1080/01402382.2014.887881>
- Toshkov, D., & Romeijn, J. (2021). How to estimate the policy preferences of party supporters: Disaggregating data from voting advice applications versus modeling survey responses. *Electoral Studies*, 74, Article 102403. <https://doi.org/10.1016/j.electstud.2021.102403>
- 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 Elsas, E. J., Miltenburg, E. M., & van der Meer, T. W. G. (2016). If I recall correctly. An event history analysis of forgetting and recollecting past voting behavior. *Journal of Elections, Public Opinion and Parties*, 26(3), 253–272. <https://doi.org/10.1080/17457289.2016.1150286>
- Walgrave, S., van Aelst, P., & Nuytemans, M. (2008). 'Do the vote test': The electoral effects of a popular vote advice application at the 2004 Belgian elections. *Acta Politica*, 43(1), 50–70. <https://doi.org/10.1057/palgrave.ap.5500209>
- Werner, A. (2020). Representation in Western Europe: Connecting party-voter congruence and party goals. *The British Journal of Politics and International Relations*, 22(1), 122–142. <https://doi.org/10.1177/1369148119873102>

About the Author



Mattia Gatti is a civil servant and PhD student in politics at LUISS University, in Rome. He was awarded a bachelor's degree in political science and a master's degree in government and policies (cum laude) at LUISS. He has served as teaching assistant in research methods for public policy and data analysis for social sciences. Contributor of the Centro Italiano Studi Elettorali (CISE) since 2020, his research interests focus on political representation, party behaviour, public opinion, and local electoral politics.



POLITICS AND GOVERNANCE

ISSN: 2183-2463

Politics and Governance is an international, peer-reviewed open access journal that publishes significant and cutting-edge research drawn from all areas of political science.

Its central aim is thereby to enhance the broad scholarly understanding of the range of contemporary political and governing processes, and impact upon of states, political entities, international organisations, communities, societies and individuals, at international, regional, national and local levels.



www.cogitatiopress.com/politicsandgovernance