Do Voting Advice Applications Have Lasting Causal Effects? Evidence from Concurrent Survey and Field Experiments

Micha Germann* Iulia Cioroianu† Fernando Mendez‡
Constantinos Djouvas§

Abstract

Voting Advice Applications (VAAs) are used by millions of voters, yet their causal impacts on political knowledge, preferences, and behavior remain unclear. We present new evidence from two complementary experiments conducted during the 2024 UK General Election: (1) a panel survey experiment with a novel design that provides access to participants' VAA advice, increases compliance with treatment assignment, and allows us to track outcomes over time; (2) a concurrent field experiment embedded in a real-world VAA that allows us to assess the ecological validity of our findings. We find that VAA usage leads to durable changes in voters' knowledge of how close they are to parties on policy issues, shifts party preferences in line with the advice received, and affects the vote choices of undecided voters. These results suggest that VAAs play an important role in election campaigns. However, we find no evidence that VAAs increase electoral participation.

Word count: 9,700

Key words: Voting advice applications, political knowledge, party preferences, electoral participation, randomized experiment

*Department of Politics, Languages & International Studies, University of Bath. m.germann@bath.ac.uk
†Department of Politics, Languages & International Studies, University of Bath.
‡Center for Democracy Studies, University of Zurich.
§Department of Communication and Internet Studies, Cyprus University of Technology.

Introduction[1]

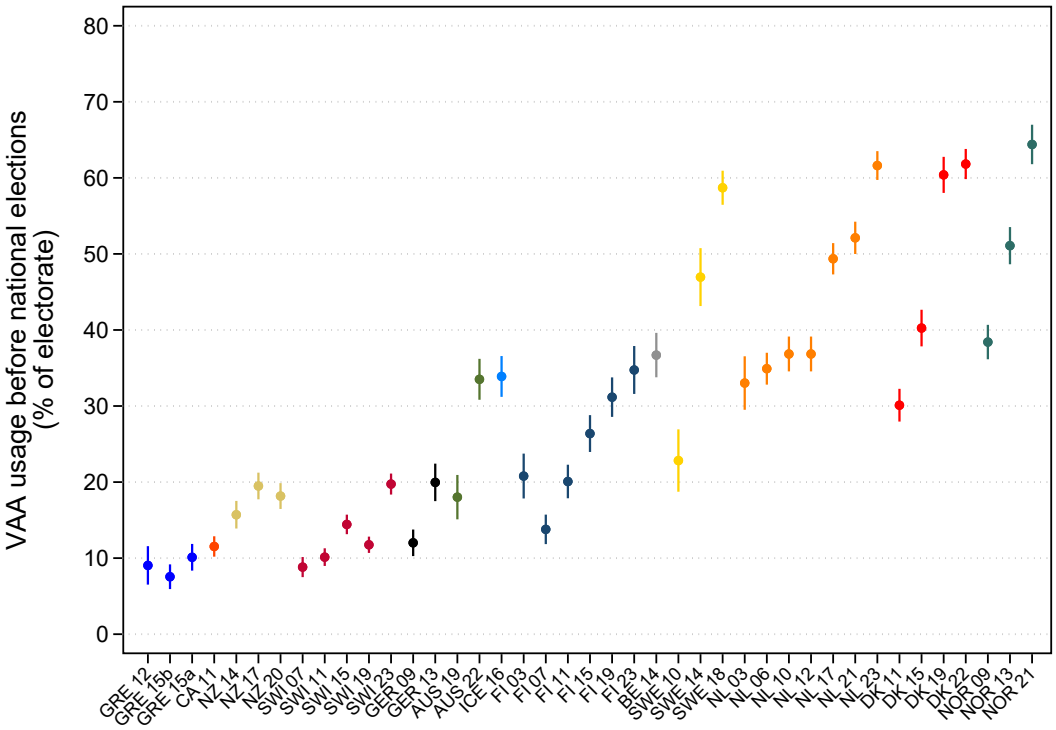
Voting Advice Applications (VAAs) are online tools that allow voters to quickly gather personalized information on how close they are to political parties or candidates on policy issues. Voters access VAAs by visiting a dedicated website where they are asked to indicate their opinions on a range of policy issues. The VAA then compares their positions with those of parties or candidates, typically coded on the basis of expert surveys, roll-call vote data, or elite surveys. Finally, users are presented with a graphical display, such as a ranked bar chart, showing how close different parties or candidates are to them on policy issues.

Initially developed in the Netherlands (De Graaf 2010), VAAs are now regularly made available before major elections in countries around the world and have achieved significant popularity. In Figure 1, we show nationally representative estimates of the share of voters who turned to VAAs before national elections in 13 countries. VAA usage has generally increased over time and more than half of voters turned to VAAs before recent elections in the Netherlands, Denmark, Norway and Sweden. Meanwhile, about every third voter did so in Belgium, Finland, and Iceland; every fifth in Germany, Switzerland, Australia, and New Zealand; and around one in ten in Canada and Greece. Similar survey data is not available from other countries, but VAAs have also been made available in countries such as France, Italy, Romania, Spain, Portugal, the United States, or the UK.

Most VAAs are championed by governmental or non-governmental organizations, typically with the goal of increasing voters' understanding of politics and promoting electoral participation. For example, the most popular VAA in Germany – *Wahl-o-Mat* – is maintained and disseminated by the Federal Agency for Civic Education, a government agency dedicated to citizen education and the promotion of democracy (Marschall and Schultze 2012). In other countries, such as Belgium, Denmark, or Finland, popular VAAs are dis-

[1] *Acknowledgements:* We would like to thank Jonathan Wheatley, Matthew Wall, Susan Banducci, Ailsa Henderson, Fraser McMillan, Roula Nezi, Ed Poole, Vasiliki Triga, and Richard Wyn Jones for their help with the development of *WhoGetsMyVote*. For excellent research assistance, we would like to thank Helen Bramah, Louis Bromfield, Keiron Burgess, Lucija Cepo, Sean Garrett, Tom Gibson, Danielle Joyce, Ellen Kear, Bella Longdon, Megan Salter, Ryan Whitehead, and Eva Whyte. Earlier versions of this paper were presented at workshops and research seminars at the University of Bath, University of Southampton, and University of Zurich. This research was supported by the Economic and Social Research Council (grant numbers: ESRCIAA-FF018 and ES/W000598/1). Study 1 has been approved by the University of Bath's Data & Digital Science Research Ethics Committee (reference number: 4749-4847). Study 2 has been approved by Oxford Brookes University's Research Ethics Committee (reference number: L19183).

Figure 1: VAA usage in 13 democracies



Note: The spikes represent 95% confidence intervals. Information on data sources, question wordings, and weighting can be found in Online Appendix §1.

seminated by national newspapers or public broadcasters, and feature in major TV shows (Walgrave, van Aelst, and Nuytemans 2008). Political scientists are frequently involved in the design of VAAs, and sometimes themselves run VAAs. For instance, political scientists based at the European University Institute have provided a VAA at every European election since 2009 (Reiljan et al. 2020). Finally, VAAs have even been used as part of international democracy-promoting initiatives. For example, in Bosnia the OSCE has collaborated with a local NGO to launch a VAA called *Glasometar* in an effort to help voters make informed decisions and promote electoral participation. However, whether VAAs actually achieve the hopes and expectations associated with them remains unclear.

A large literature has studied the effects of VAA usage on political knowledge, electoral preferences, and turnout. Despite this, it remains unclear whether VAA usage has any durable effects. Broadly speaking, there are two types of VAA studies – observational studies and randomized experiments – and these often report diametrically opposed results. Most observational studies report evidence in favor of substantially sized VAA effects on

political knowledge (e.g., Schultze 2014), electoral preferences (e.g., Alvarez et al. 2014; Kleinnijenhuis et al. 2019; Tromborg and Albertsen 2023), and electoral participation (e.g., Garzia, de Angelis, and Pianzola 2014; Germann and Gemenis 2019; Gemenis and Rosema 2014; Kruikemeier et al. 2014). Yet, most experimental studies have been unable to replicate these effects (e.g., Enyedi 2016; Mahéo 2016, 2017; Munzert et al. 2020; Pianzola et al. 2019; Vassil 2011). The reasons for this discrepancy remain unclear. One possible explanation is that observational studies are biased since they do not adequately account for self-selection into VAA usage. Yet, experimental studies may also be biased towards the null due to limitations of their experimental design (Germann, Mendez, and Gemenis 2023; Munzert and Ramirez-Ruiz 2021).

We revisit the effects of VAA usage on political knowledge, party preferences, and electoral participation using a new experimental design that addresses the limitations of previous experimental studies (study 1). Similar to most previous studies, our experimental manipulation consists of a prompt inviting participants in an online survey to use a VAA. However, unlike previous studies, we do not merely encourage VAA usage, but directly route participants to an external VAA website and then back to our survey. Since treated participants can only complete the survey (and receive the payment) by completing the VAA, our set-up significantly increases compliance. Furthermore, our new design also guarantees access to the VAA advice seen by treated participants, and the VAA advice that control participants would have seen had they been assigned to the treatment group. We facilitate this by tracking the behavior of the treatment group on the VAA platform and by asking the control group the same set of policy questions presented in the VAA, but without exposing them to the VAA advice. By contrast, most previous experimental studies lacked access to participants' VAA advice, which makes it difficult to effectively test VAA effects.

We embedded our experimental manipulation in a panel survey administered to more than 3,000 British voters before and shortly after the 2024 UK General Election (study 1). The panel structure allows us to track the causal effects of VAA usage over time.The UK represents a good context for our study since VAAs have not achieved the widespread popularity seen in some other countries, which reduces the likelihood of contamination in the control group and thus improves our ability to estimate causal effects (Germann and

Gemenis 2019).

Contrary to most previous experimental studies, we find clear evidence that VAA usage increases voters' ability to identify how close they are to parties on policy issues, and that VAA users tend to align their party preferences with the advice they receive. These effects are strongest in the immediate aftermath of VAA usage and decrease subsequently. However, VAA effects on proximity perceptions and party preferences remain substantially sized after the election and, importantly, we find that VAA usage affected the actual choices of undecided voters at the ballot box. Additional analyses suggest that these results are driven by our ability to capture participants' VAA advice: when we replicate the default approach in previous studies and estimate the effect of VAA usage on vote switching (i.e., change in party preferences over the course of a campaign), we find no evidence for any VAA effects.

Similar to the majority of existing experimental studies, study 1 artificially induces VAA usage and thus raises questions about the generalizability of causal effects from the experimental setting to real-world behavior. To investigate the ecological validity of our findings, study 2 implements a concurrent field experiment relying on a previously proposed, yet rarely used, experimental design that randomizes the time when real-world VAA users are asked about relevant outcome questions – either before or after exposure to the VAA advice (Garry et al. 2019; Germann, Mendez, and Gemenis 2023). This allows us to compare the immediate causal effects of exposure to VAA advice in a real-world setting with the effects of artificially induced VAA usage. We find evidence that the effects of artificially induced VAA usage are likely to generalize to the real world.

Overall, our results suggest that methodological limitations may have led to overly pessimistic conclusions about the effectiveness of VAAs in previous experimental research. VAAs help citizens make more informed voting decisions and shape party preferences. As we argue below, these effects are likely to be stronger in countries where VAAs are most popular since the latter, unlike the UK, tend to have proportional electoral systems with fewer incentives for tactical voting. Finally, we find little evidence for an effect of VAA usage on electoral participation in either of our experimental studies. This finding is consistent with the majority of existing experimental studies, and casts doubt on observational findings suggesting that VAAs are an effective mechanism to help reverse declining turnout rates.

Methodological Challenges in VAA Research

Normatively, it is often seen as desirable if voters base their electoral choices on policy considerations (Achen and Bartels 2017). However, empirical studies indicate that many citizens lack a clear understanding of the policy positions held by parties and candidates (Adams, Ezrow, and Somer-Topcu 2011; Grand and Tiemann 2013; Merrill, Grofman, and Adams 2001). At least in part, this knowledge gap is likely attributable to the high cognitive and informational costs associated with acquiring information on policy stances through traditional channels, such as newspapers, broadcast media, or party manifestos. Voters often lack the incentives to incur these costs and many therefore cast votes that do not align with their underlying policy preferences (Lau et al. 2014; Rogers 2017; Wheatley and Germann 2024). VAAs aim to alter this equation by providing carefully crafted *personalized* information on voters' policy congruence with parties or candidates (Walgrave, van Aelst, and Nuytemans 2008). However, due to a series of methodological challenges, it remains unclear to what extent voters internalize the information shown by VAAs, and whether it affects their electoral preferences and participation.

The effects of VAA usage have been widely studied using a variety of different research designs. Especially in early studies, a popular approach has been to turn to election surveys and estimate correlations between VAA usage and political knowledge (Schultze 2014), voting preferences (Andreadis and Wall 2014; Walgrave, van Aelst, and Nuytemans 2008), or electoral participation (Garzia, de Angelis, and Pianzola 2014; Gemenis and Rosema 2014; Germann and Gemenis 2019; Kruikemeier et al. 2014; Marschall and Schultze 2012). Most of these studies report substantially sized VAA effects; however, correlations in cross-sectional election studies assume selection on observables, which is a strong assumption (Pianzola 2014a). More recent observational studies have therefore turned to more sophisticated designs, including panel (Alvarez et al. 2014; Benesch et al. 2023; Kamoen et al. 2015; Kleinnijenhuis et al. 2019; Heinsohn et al. 2019; Tromborg and Albertsen 2023) and instrumental variables designs (Garzia, de Angelis, and Pianzola 2014; Pianzola 2014b). Most of these more advanced studies continued to find evidence for VAA effects; however, causal identification remains linked to strong assumptions in both panel studies and instrumental

variable approaches.

In response to endogeneity concerns, VAA research has increasingly turned to randomized experiments and, casting significant doubt on observational estimates, most experimental studies have not been able to replicate the results of observational studies (Munzert and Ramirez-Ruiz 2021). However, there is important variation according to the design of experimental studies. Most previous studies have drawn on the encouragement design, which involves the random assignment of a verbal appeal to use a VAA (sometimes combined with a financial incentive). Many of these studies have reported null results (Enyedi 2016; Garzia, Trechsel, and de Angelis 2017; Mahéo 2016, 2017; Munzert et al. 2020; Pianzola et al. 2019; Vassil 2011). Yet, two other recent studies have drawn on an alternative experimental design – the timing design – which randomizes whether real-world VAA users are asked about outcomes of interest before or after they are exposed to the VAA advice. These studies found consistent evidence for a VAA effect on party preferences (the only outcome they studied) (Garry et al. 2019; Germann, Mendez, and Gemenis 2023). Unfortunately, though, both the encouragement and the timing design have limitations that could compromise the validity of their conclusions. We discuss each in turn.

Initially proposed by Vassil (2011), the encouragement design exogeneizes VAA usage and therefore makes it possible to disentangle the effects of VAA usage from potential confounders. A second critical advantage of the encouragement design is that it facilitates the measurement of the effects of VAA usage after several days or even weeks, and, therefore, a longer and more substantively meaningful time span compared to timing studies (see below). However, the encouragement design also has several significant limitations.

First, the encouragement design does not provide access to the voting advice seen by VAA users. This is problematic since we would expect that VAA usage shifts voters' candidate and party evaluations in line with the advice they receive (Germann et al. 2024; Wall, Krouwel, and Vitiello 2014). Lack of access to the VAA advice therefore prevents researchers from effectively testing VAA effects.[2] In an attempt to circumvent this problem, encouragement studies tend not to estimate the effects of VAA usage on party preferences as such, but on

[2] Lack of access to the VAA advice is a problem also in some observational studies (though see Alvarez et al. 2014; Tromborg and Albertsen 2023; Wall, Krouwel, and Vitiello 2014).

the rate of vote switching (Mahéo 2016; Munzert et al. 2020; Pianzola et al. 2019; Vassil 2011).[3] However, depending on the advice that users receive, VAAs may both lead to preference change *and* preference stabilization. Hence, the average effect of VAA usage on vote switching could be positive, negative, or zero (Tromborg and Albertsen 2023). In other words, a null effect on vote switching does not preclude an effect on party preferences.

A second problem of many existing encouragement studies is low compliance with treatment assignment (Garzia, Trechsel, and de Angelis 2017; Germann and Gemenis 2019). Participants who are encouraged to use a VAA may not use the VAA despite researcher encouragement ("never takers"). Furthermore, participants may use VAAs independent of researcher encouragement ("always takers"), a problem that tends to be particularly acute in countries where VAAs are very popular. For example, in a recent German study, only 65% of participants encouraged to use a VAA actually used a VAA while 29% of participants did so even though they did not receive the encouragement (Munzert and Ramirez-Ruiz 2021). Similarly, in a recent Swiss study 83% of participants assigned to the treatment group used a VAA, but so did 69% of participants assigned to the control group (Pianzola et al. 2019). Non-compliance of this magnitude is likely to lead to severe downward biases in intention-to-treat analyses, and may also bias the results of selection models designed to estimate complier effects (Keele and Morgan 2016; Murray 2006). Often, these problems are compounded by small samples and consequently low statistical power (Munzert and Ramirez-Ruiz 2021).

Finally, encouragement studies estimate the effects of artificially induced VAA usage, which may not generalize well to real-world VAA usage. Actual voters do not use VAAs because they are encouraged to do so (and potentially paid) by a researcher, but because they themselves decide to use a VAA. Real-world VAA users may therefore have greater intrinsic motivation, which could lead to larger effects. At the same time, effects in the real world may also be smaller. For example, the prospect of a financial reward could increase attention levels in encouragement studies. Ultimately, it remains unclear how the effects of artificially induced VAA usage compare to "organic" VAA usage.

[3] Another, though less frequent, approach has been to ask participants to recall which party or candidate they were recommended (Enyedi 2016). However, recall measures may be subject to cognitive biases and therefore bias causal estimates.

The timing design addresses these limitations, albeit at a significant cost (Garry et al. 2019; Germann, Mendez, and Gemenis 2023). Unlike the encouragement design, timing studies embed the experimental manipulation directly in a live VAA. This allows researchers to estimate VAA effects among real VAA users with high compliance rates and high statistical power (given VAAs' broad usage). Furthermore, the timing design provides access to the VAA advice seen by subjects in the treatment group *and* the VAA advice that the control group receives subsequently (but has not seen yet). This makes it straightforward to test whether candidate or party evaluations shift in line with the advice using multiplicative interactions (Germann, Mendez, and Gemenis 2023). However, the timing design also has a critical disadvantage. What is randomized is not VAA usage, but whether VAA users have already been exposed to the VAA advice. Since both the treatment and the control group ultimately receive the advice, the timing design can only identify the *immediate* effects of exposure to VAA advice. While timing studies present causal evidence that VAA users update their party preferences after VAA usage, it therefore remains unclear how durable these effects are.

We contribute to the debate on the effects of VAA usage using a two-pronged empirical strategy. First, we introduce a new experimental design that, similar to the encouragement design, artificially induces VAA usage, but does so in a way that significantly increases compliance with treatment assignment and gives us access to the VAA advice for both the treatment and the control group. To mitigate concerns about statistical power, we field this experiment to a sample that is considerably larger compared to previous encouragement studies, and we focus on a country (the UK) where VAAs are not massively popular, thus reducing the risk of contamination in the control group. To gain insight into the durability of VAA effects, we repeatedly survey our participants after the experimental manipulation, allowing us to compare the effects of VAA usage immediately after usage, a few days after usage, and after the election. Second, we conduct a concurrent field experiment using the timing design. For the first time, this lets us directly compare the immediate effects of exposure to VAA advice in the real world with the effects of artificially induced VAA usage. Thus, we can gain insight into the generalizability of the effects of artificially induced VAA usage.

Theoretical Expectations

We assess the effects of VAAs on all three main outcomes studied in previous literature. First, in keeping with existing literature we expect that VAA usage increases political knowledge (Heinsohn et al. 2019; Munzert et al. 2020; Schultze 2014). However, in deviation from previous studies we do not test the effects of VAA usage on voters' ability to correctly identify the policy stances of parties, but on voters' ability to identify how close different parties are to them on policy issues. Fundamentally, the central output of VAAs is not information on the policy stances of parties, but information on how close users are to parties (or, depending on the VAA, candidates) on policy issues. While many VAAs also provide information on policy stances, this is not what VAA users are presented with first, and many VAA users never even access this additional information (Manavopoulos, Mendez, and Djouvas 2020). Hence, the primary knowledge gain from VAAs is likely to come in the form of increased awareness of policy proximity to parties or candidates. Given that the VAA we study below shows proximity estimates for parties and not individual candidates, this leads to our first hypothesis:

> *H1: VAA users tend to align their perceived proximity to parties on policy issues with the advice they receive.*

Second, we expect that voters update their electoral preferences after VAA usage. Existing research suggests that voting preferences are influenced by a myriad of factors, including affective attachment to parties (Cohen 2003), the country's economic performance (Lewis-Beck and Stegmaier 2018), valence factors (Green 2007), and tactical considerations (Riera 2015). As suggested by issue voting theory, the policy positions of political parties or candidates represent another important factor (Enelow and Hinich 1984). Indeed, survey evidence suggests that proximity on policy issues is one of, if not the most important consideration for a majority of voters (e.g., Fieldhouse et al. 2022; Borelli 2023). Yet, despite this existing research suggests that voters' electoral preferences often do not match their policy preferences (Hanretty, Mellon, and English 2021; Lau et al. 2014; Rogers 2017; Wheatley and Germann 2024). Furthermore, research indicates that electoral preferences are often uncertain and amenable to new information (Alvarez 1998; Lau and Redlawsk 2006). VAAs

provide voters with carefully curated, personalized information on their policy proximity to parties or candidates. To the extent that voters see issues as an important consideration, they may therefore update their electoral preferences after VAA usage (Germann, Mendez, and Gemenis 2023; Munzert et al. 2020; Pianzola et al. 2019; Walgrave, van Aelst, and Nuytemans 2008). This leads to our second hypothesis:

H2: VAA users tend to align their party preferences with the advice they receive.

Third, we expect that VAA usage increases electoral participation. Existing literature points to two main causal mechanisms linking VAA usage to increased turnout. First, VAAs provide easily accessible information on issue congruence and therefore significantly reduce the information costs associated with electoral participation (Garzia, Trechsel, and de Angelis 2017; Gemenis and Rosema 2014; Mahéo 2017). Second, VAAs highlight that there are important differences between the policy platforms of different parties or candidates, and often show users that they are much more closely aligned with some parties or candidates compared to others. Hence, VAA usage may increase the instrumental benefits associated with voting (Germann and Gemenis 2019; Marschall and Schultze 2012).

H3: VAA usage increases the probability of electoral participation.

Finally, we consider how durable the effects of VAAs are. Previous research suggests that voters often forget information they acquire during election campaigns, and this likely extends to VAAs (Dalager 1996; Lodge, Steenbergen, and Brau 1995). One common explanation for why people forget information is memory decay: memory traces fade over time and become increasingly less available for retrieval (Brown 1958; Joslyn 2003). Another common explanation is interfering information (Fawcett and Hulbert 2020; Wixted 2004). As the interval between VAA use and election day lengthens, other information voters acquire during election campaigns may increasingly crowd out the information learned from VAAs (Miller 2013). However, while the effects of VAA usage are therefore likely to decrease as time passes, we still expect that VAA effects remain visible over the course of an election campaign. VAAs tend to be released in relatively close temporal proximity to elections (i.e., several days or weeks before election day), and the information they show is unique. This

makes it likely that VAA users, or at least some VAA users, retain the VAA advice over the course of an election campaign.

> *H4: The effects of VAA usage decrease but remain visible over the course of an election campaign.*

Study 1

We test our hypotheses using a randomized experiment embedded in a four-wave panel survey conducted before and shortly after the 2024 UK General Election. Our experimental design facilitates near-universal treatment uptake, lets us estimate VAA effects conditional on the advice received, and allows us to trace how VAA effects change over time. We pre-registered our design, hypotheses, and analysis plan with AsPredicted.[4] Minor departures from the pre-registration are reported in Online Appendix §2.1.

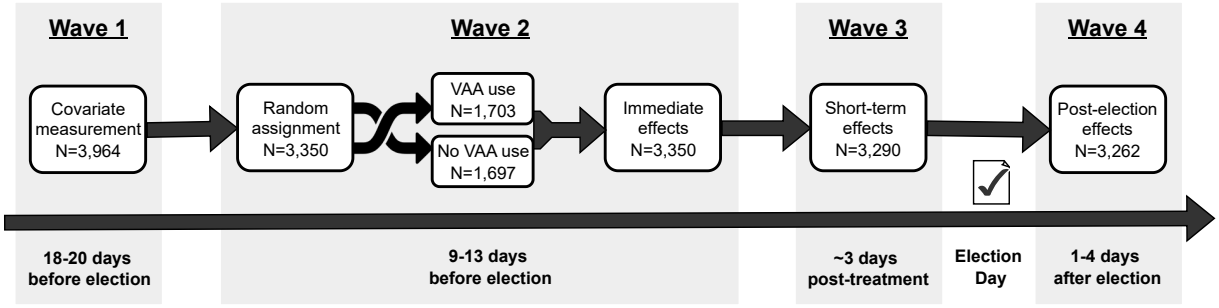
Experimental Design

We recruited participants from Prolific, a research-oriented online access panel. Participants needed to be eligible to vote in the 2024 UK General Election and be resident in England, Scotland, or Wales. We initiated the first panel wave ($N = 3,964$) 20 days before the election to collect pre-treatment information on socio-demograhpics and our outcome variables. Participants were informed that they will be invited to three more surveys over the coming weeks, and they were promised a bonus payment if they participate in all surveys. We had survey quotas in place to render the sample representative of the British adult population in terms of sex, age, and region residence. Online Appendix §2.2 shows sample descriptives.

The second wave was initiated about a week after the first and included our experimental manipulation (see Figure 2). Half of wave 2 participants were randomly assigned a prompt asking them to complete a VAA called *WhoGetsMyVote* (see Online Appendix §2.3 for the complete prompt). The prompt informed participants assigned to the treatment group that *WhoGetsMyVote* is an online voter information tool that will show them how close

[4] See AsPredicted #179371 for the turnout analysis and AsPredicted #179379 for the perceived proximity and party preferences analyses.

Figure 2: Overview of study 1

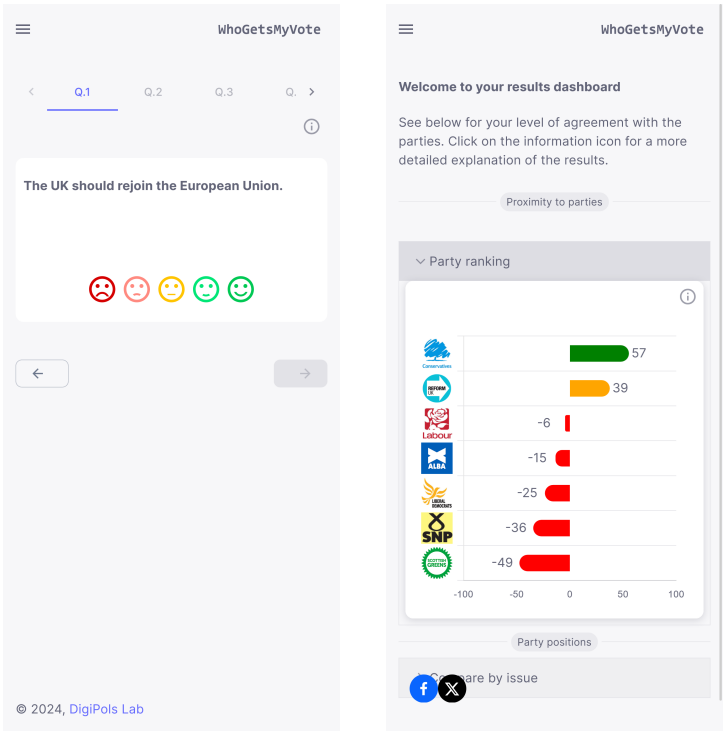


political parties are to their preferred policy positions. Participants were told that they must visit an external website and will then return to the survey platform. Finally, they were informed that they would only be able to proceed with the survey if they entered a completion code, which will appear after 30 seconds on the VAA results screen in the form of a pop-up message. This creates strong incentives for participants to comply and complete the VAA. Prolific requires that participants receive a financial reward for every survey they take, but receipt of that reward is tied to survey completion. Hence, almost all of the 1,703 participants assigned to the treatment condition (99.7%) completed the VAA.[5] Meanwhile, the control group was not asked to complete *WhoGetsMyVote* and never left the survey platform.

The design of *WhoGetsMyVote* closely followed that of other VAAs. First, VAA users were asked to select their region of residence and then received information on the purpose of the application and the data protection policy. Next, VAA users were asked to provide their opinions on 25 salient policy issues using five-point Likert scales ranging from 'Completely disagree' to 'Completely agree'. Finally, the application compared users' policy preferences with expert-coded party positions, and then showed the results in the form of a bar chart ranking the different parties from closest to most-distant across all 25 policy issues. The number of parties included in this ranking included five in England, six in Wales, and seven in Scotland. Proximity between users and parties on policy issues was calculated using the City Block algorithm (Mendez 2012). Figure 3 shows an example policy question and VAA

[5] A small number of participants assigned to the treatment group (6%) did not comply initially, but completed the VAA after being reinvited. 4% of participants completed the VAA but not the outcome questions, presumably because they forgot the completion code. We reinvited those respondents to answer the outcome questions and thus claim their financial reward.

Figure 3: Example policy question and VAA advice (*WhoGetsMyVote*)



(a) Policy question (b) VAA party ranking

Note: To optimize the user experience on small screens, *WhoGetsMyVote* users were asked to indicate their agreement to policy proposals using a smiley rating system. It was clearly explained that the different smileys represent different levels of agreement, ranging from complete disagreement (dark red, unhappy smiley) to complete agreement (dark green, happy smiley).

result. Online Appendix §2.4 contains additional details on the design of *WhoGetsMyVote*, including how we coded party positions and the wordings of all policy questions.

Critically, our design allows us to access the VAA advice seen by treated participants *and* to establish the VAA advice that control participants would have seen had they been assigned to the treatment group. For the treatment group, we facilitated this by personalizing the weblink that participants used to visit *WhoGetsMyVote* with a unique identifier. This allows us to match the VAA records of treated participants with our survey records and, therefore, to recreate the advice they saw. Meanwhile, we asked the control group the same 25 policy questions that the treatment group completed as part of the VAA. Hence, we can establish the advice control participants would have received had they been assigned to the treatment group. To prevent survey mode effects, we used the same smiley rating system used in the VAA (see Figure 3). As we would expect under random assignment, the treatment and control groups are indistinguishable in terms of their socio-demographic profile and the

VAA advice they received or would have received (see Online Appendix §2.5).

After the experimental manipulation, both the treatment and control groups were asked the same set of outcome questions (see below). All wave 2 participants were then re-invited twice to answer a similar set of outcome questions. Wave 3 was administered 3.2 days after the treatment wave on average and wave 4 the day after the election (12.8 days after wave 2 on average). As is common in panel studies, there was some attrition between panel waves. However, the level of attrition was relatively small, especially from wave 2 onward. Specifically, out of 3,964 wave 1 participants, 3,350 returned and completed wave 2, or about 85%. Almost all wave 2 participants then returned to complete wave 3 (98%) as well as wave 4 (97%). Panel attrition is comparable in the two experimental groups (see Online Appendix §2.6).

Outcomes

We assess the effects of VAA usage on participants' (1) perceived proximity to parties on policy issues, (2) party preferences, and (3) electoral participation. First, we measured participants' perceived policy proximity by asking them to rate, on a scale from 0 to 10, how close the parties included in the VAA advice are to them on policy issues.

Second, we measure party preferences by asking participants to rate, on a scale from 0 to 10, the likelihood that they would ever vote for the parties included in the VAA advice. In addition, we asked participants which party they intend to vote for (waves 2 and 3) or which party they actually voted for (wave 4) in the 2024 UK General Election. These two measures have complementary strengths. On the one hand, the propensity to vote questions allow us to establish the effects of VAA usage on participants' complete set of party preferences. On the other hand, the binary vote intention and vote choice questions allow us to test whether VAA usage has consequences for voters' intended and actual behavior at the ballot box.

Finally, we measure electoral participation in two ways. First, we asked participants to rate, on a scale from 0 to 10, how likely it is that they participate in elections. Second, we asked participants whether they intend to participate (waves 2 and 3) or actually participated (wave 4) in the 2024 UK General Election. Online Appendix §2.7 contains the exact wordings of all outcome questions and summary statistics.

Estimation

Our hypotheses make different types of predictions depending on the outcome variable. First, we expect a simple average increase in the probability of electoral participation after VAA usage. This allows us to test the effects of VAA usage on turnout using the following, straightforward regression set-up:

$$Y_{iw} = \alpha + \beta_1 D_i + \beta_2 X_i + \varepsilon_{iw} \tag{1}$$

where Y_{iw} represents participant i's actual or intended electoral participation, as declared in panel wave w, D_i whether or not participant i was assigned to the treatment group, X_i a vector of control variables, and ε_{iw} an error term. The list of control variables includes a standard set of socio-demographics as well as pre-treatment information on a participant's turnout intention (see Online Appendix §2.8 for variable definitions and summary statistics). Controlling for pre-treatment outcomes increases the precision of causal estimates (Clifford, Sheagley, and Piston 2021). We estimate this model using standard ordinary least squares (OLS) regression.

Second, we expect that the effects of VAA usage on perceived policy proximity and party preferences depend on the VAA advice that users receive. To capture the conditional nature of these expectations, we follow Garry et al. (2019) and Germann, Mendez, and Gemenis (2023) and estimate linear multiplicative interaction models of the following form:

$$Y_{ijw} = \alpha + \beta_1 D_i + \beta_2 Z_{ij} + \beta_3 D_i Z_{ij} + \beta_4 X_{ij} + \gamma_j + \varepsilon_{ijw} \tag{2}$$

where Y_{ijw} represents participant i's perception of proximity to party j in wave w, or, alternatively, participant i's preference for party j in wave w. As before, D_i represents participant i's treatment status, while Z_{ij} represents the rank party j had, or would have had, in participant i's VAA advice. More specifically, Z_{ij} represents a vector of binary variables indicating whether a given party was, or would have been, ranked first or not, second or not, third or not, etc. in the VAA advice. Since only treated participants were actually exposed to the VAA advice the interaction between treatment status and the VAA party ranking ($D_i Z_{ij}$) reflects the effect of VAA usage. We dichotomize the VAA party

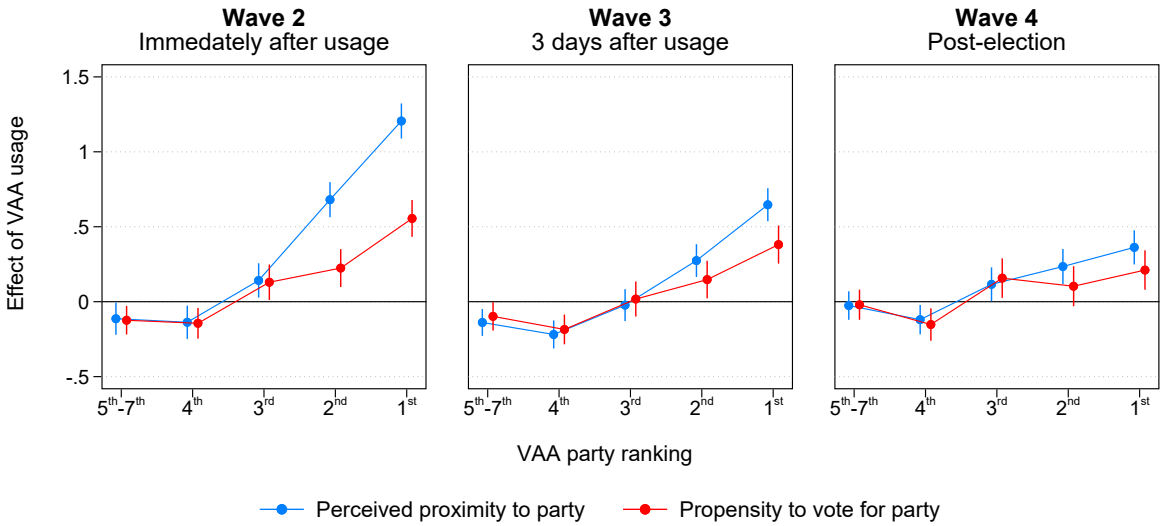
ranking so as to flexibly estimate treatment effects for different party ranks. As before, we control for pre-treatment measures of our outcome variables and socio-demographics (X_{ij}) and use OLS regression for estimation. However, contrary to the turnout analysis we now also include party fixed effects (γ_j) to account for unobserved party-level heterogeneity. Furthermore, since participants rated multiple parties and are therefore observed more than once, we cluster standard errors at the participant level.

Results

We begin by discussing the effects of VAA usage on the perceived proximity to parties on policy issues (see the blue estimates in Figure 4). Consistent with theoretical expectations, we find that VAA users updated their proximity perceptions after VAA usage, and that this effect is largest immediately after VAA usage. Specifically, a VAA user rated the top-ranked party about 1.2 points closer on policy issues compared to a non-user immediately after usage ($p < 0.001$), the second-ranked party about 0.7 points closer ($p < 0.001$), and the third-ranked party about 0.1 points closer ($p < 0.05$). These represent sizable increases of approximately 5% to almost 40% of a standard deviation. Meanwhile, treated participants saw the parties that are ranked lowest by the VAA as slightly less close after VAA usage (about 0.1 points; $p < 0.05$). Hence, VAA usage can both increase and decrease perceived proximity on policy issues. However, it is notable that the negative effects are comparatively small. This may be because voters find it easier to identify parties they disagree with and harder to adjudicate between parties they have some agreement with. Another reason could be that VAA users are more focused on the top part of the VAA party ranking, which is the most relevant for electoral decisions.

Over time, the effects of VAA usage on perceived policy proximity decrease in size, but as predicted they remain visible after the election. For example, the boost experienced by the top-ranked party decreases from +1.2 points immediately after usage to +0.6 points (or +20% of a standard deviation) three days after VAA usage ($p < 0.001$), and to +0.4 points (or +11% of a standard deviation) after the election ($p < 0.001$). This points to significant effect dissipation. Still, these results imply that a VAA user rated the top-ranked party as substantially closer on policy issues almost two weeks after VAA usage compared to a

Figure 4: Effects of VAA usage on perceptions of policy proximity and the propensity to vote for parties



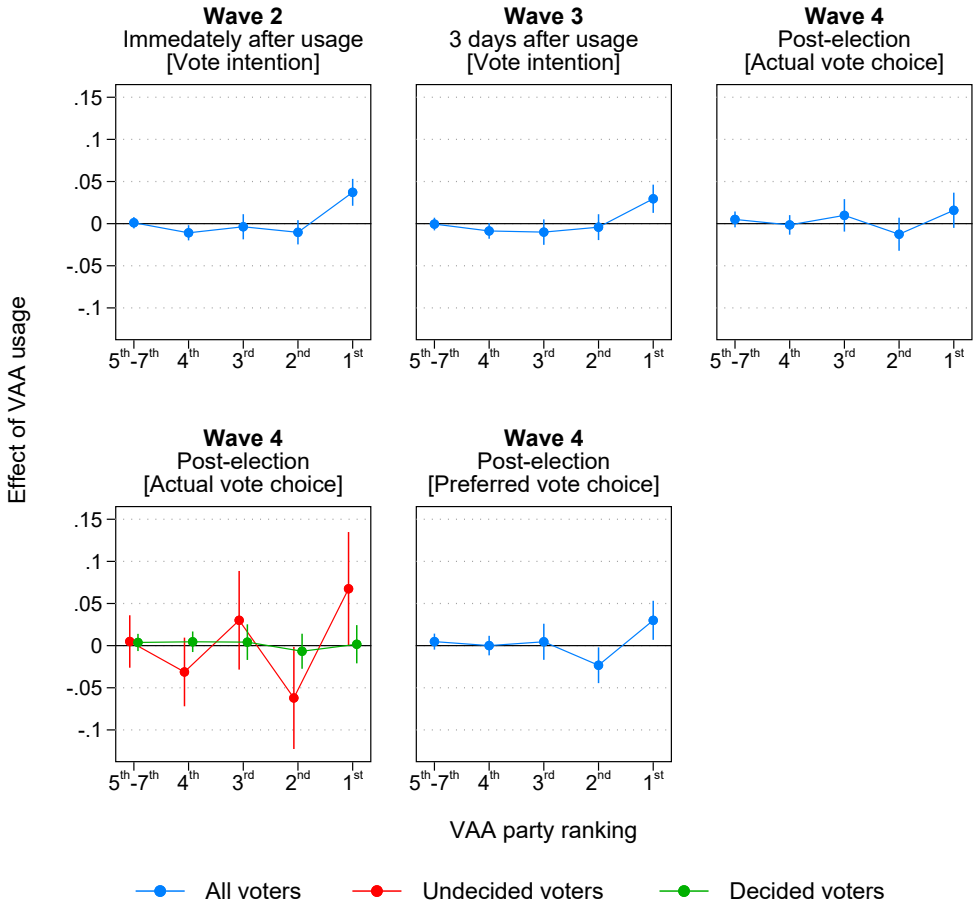
Note: All effects are conditional on the rank of a party in the VAA. Only the treatment group was actually exposed to this ranking. The spikes represent 95% confidence intervals.

non-user.

The results thus far imply that VAA usage leads to durable improvements in voters' ability to identify how close they are to parties on policy issues. Next, we consider whether VAAs also affect their users' party preferences. As the red estimates in Figure 4 show, this was indeed the case: VAA users expressed a higher propensity to vote for a party if they were told that they are close to that party on policy issues, and a lower propensity if they were told that a party is far away. As before, these effects decrease over time but remain visible after the election. However, what is notable is that the effects of VAA usage on party preferences are smaller compared to the effects on perceived policy proximity. For example, a VAA user's propensity to vote for the top-ranked party increased by about half as much when compared to proximity beliefs: +0.6 points (or +16% of a standard deviation) immediately after usage ($p < 0.001$), +0.4 points (or +11% of a standard deviation) after three days ($p < 0.001$), and +0.2 (or +6% of a standard deviation) after the election ($p < 0.01$). The likely explanation is that voters do not care solely about policy proximity, but also about other factors, such as a party's performance in government, leadership, or chances of winning.

The propensity to vote questions allow us to establish the overall effect of VAA usage on party preferences. However, in most electoral settings, voters ultimately cast a vote for

Figure 5: Effects of VAA usage on vote intention, actual vote choice, and hypothetically preferred vote choice



Note: All effects are conditional on the rank of a party in the VAA. Only the treatment group was actually exposed to this ranking. The spikes represent 95% confidence intervals.

a single party. Next, we therefore consider whether VAA usage affected the intended and actual vote choices of participants. We find that VAA usage boosted the intention to vote for the top-ranked party (and only the top-ranked party) by 3.7 percentage points immediately after usage ($p < 0.001$) (see Figure 5). This effect remains similarly sized (+3 percentage points; $p < 0.001$) three days after usage; however, it decreases to +1.6 percentage points and is no longer statistically significant after the election ($p = 0.14$). There are, however, two important qualifications to this result.

First, while VAA usage had no statistically significant average effect on actual vote choice, it had a statistically significant effect on the vote choices of an important subset of voters: undecided voters. In Online Appendix §2.9, we investigate several (pre-registered) hypotheses about treatment effect heterogeneity. One of our key findings is that VAA usage

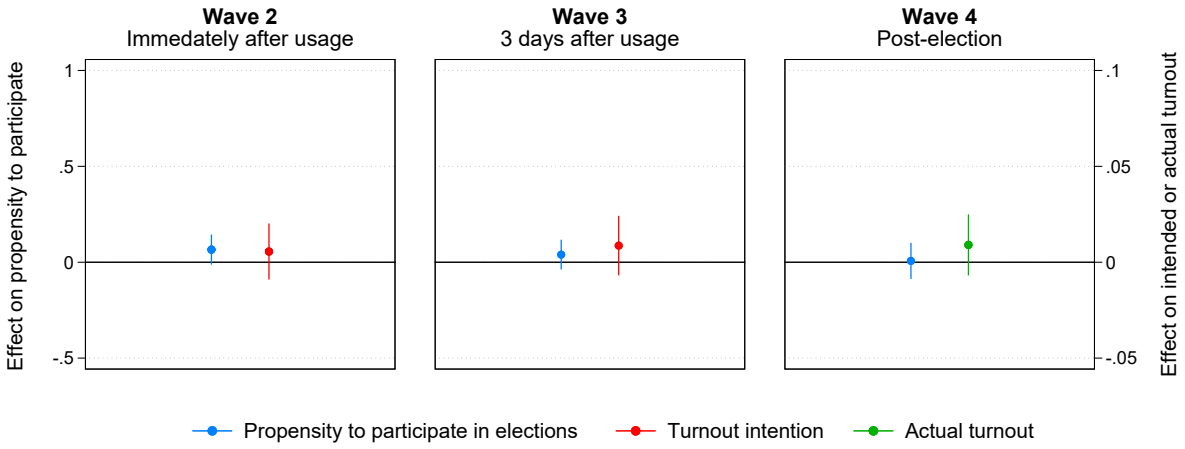
tends to have stronger effects on the party preferences of undecided voters. The bottom-left panel in Figure 5 shows a key implication: while VAA usage had practically zero effect on the vote choices of decided voters (i.e., voters who already had a clear vote intention before VAA usage), it increased the probability that voters who did not have a clear vote intention before VAA usage cast a vote for the top-ranked party by 6.7 percentage points ($p < 0.05$).[6]

Second, the UK is likely to constitute a comparatively hard case when it comes to VAAs affecting vote choice. Voters in a first-past-the-post electoral system need to consider that smaller parties may not stand a realistic chance of winning in their constituencies. Due to tactical considerations, changes in party preferences may therefore often not translate into changes in vote choice in the UK. Yet, many of the countries where VAAs are most popular have proportional electoral systems (see Figure 1), where tactical considerations are less pronounced. This would suggest that VAAs are likely to have stronger effects on vote choice where they are most popular.

While our single-country experiment does not allow us to directly test this hypothesis, we can get a basic sense of the extent to which VAA effects on vote choice are likely to increase in the absence of tactical constraints. In panel wave 4, we asked participants why they voted for the party they did, and if participants said they voted tactically, we subsequently asked them which party they would have actually preferred. Hence, we can establish the hypothetically preferred choices of tactical voters, and take these into account when estimating the effects of VAA usage. As the bottom-center panel in Figure 5) shows, the average effect of VAA usage on voting for the top-ranked party almost doubles in size, from +1.6 to +3 percentage points, when we replace the actual choices of tactical voters (about 14% of our sample) with their hypothetically preferred choices. Notably, this effect is statistically significant ($p < 0.05$). Though tentative, these findings suggest that VAAs may exert significantly stronger effects on vote choice in countries where voters are freer to express their preferences at the ballot box.

[6] In addition, the sub-group analysis suggests that VAA usage has stronger effects on the perceived proximity to parties and party preferences of younger voters, as well as voters with lower educational attainment. However, we find no evidence that VAA usage had a stronger effect on the vote choices of younger voters or voters with lower educational attainment. Interestingly, neither party identification nor political interest moderate the effects of VAA usage. The (null) result for party identification is particularly notable: according to our results, voters who identify with a party are just as likely to update their party evaluations as voters who do not identify with a party.

Figure 6: Effects of VAA usage on electoral participation

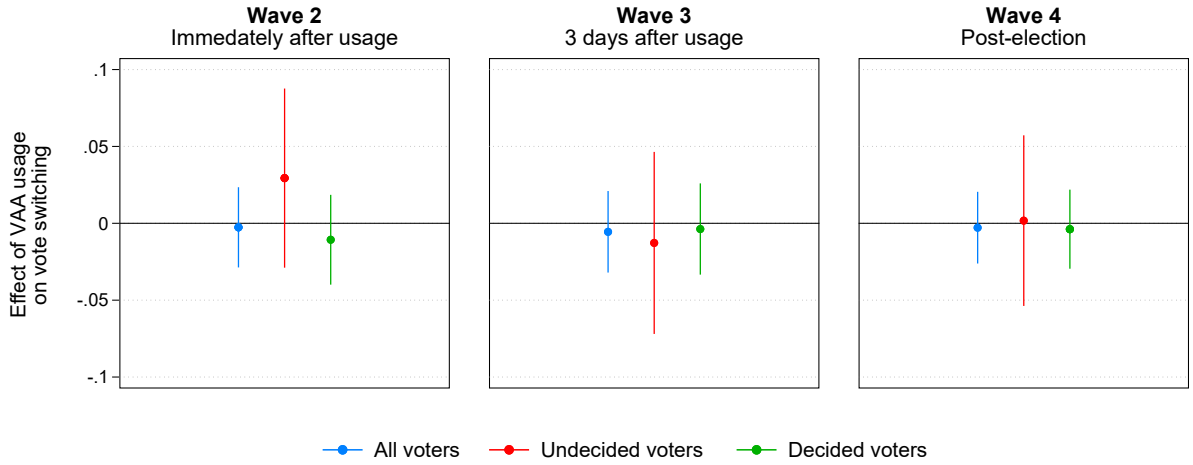


Note: The spikes represent 95% confidence intervals.

Finally, we consider the effects of VAA usage on electoral participation. Figure 6 shows the average effects of VAA usage on the general propensity to participate in elections as well as intended and actual turnout in the 2024 UK General Election. Contrary to theoretical expectations, we find no evidence for a VAA effect on electoral participation. While the point estimates are consistently positive, they never reach statistical significance. This conclusion applies irrespectively of how we measure electoral participation, and it applies in all panel waves. We also do not find an effect on electoral participation in any of the sub-groups we consider. The only partial exception is political interest: VAA users who expressed low political interest reported a higher intention to participate in the election immediately after VAA usage. However, this effect vanishes shortly afterwards (see Online Appendix §2.9).

Overall, the results of study 1 diverge from the findings of most previous experimental studies. Similar to the latter, we find no evidence for a VAA effect on electoral participation, yet contrary to previous studies, we find clear evidence that VAA usage has durable effects on voters' party evaluations. To investigate the reasons for this discrepancy, we replicate the kind of analysis that previous evaluations of VAA effects on party preferences tended to report: the average effect of VAA usage on vote switching (i.e., preference change) over the course of the campaign. As Figure 7 shows, we find no evidence for an effect on vote switching, even among undecided voters. This underlines the importance of VAA studies having access to the advice participants received (or would have received, if assigned to

Figure 7: Effects of VAA usage on vote switching



Note: We treat participants as vote switchers if their vote intention (waves 2 and 3) or actual vote choice (wave 4) differed from their vote intention in wave 1. The results are based on linear regressions including the same set of socio-demographics and political attitudes. The spikes represent 95% confidence intervals.

the control group). Despite the higher compliance rates and increased statistical power of our study, we too would have concluded that VAAs do not affect party evaluations in the absence of data on the VAA advice.

Robustness Checks

A potential concern with the intention-to-treat effects we report above is that VAAs were publicly available during the election campaign. Hence, it is possible that control participants were also exposed to VAA advice, either through the public version of our tool (see below) or through similar other tools that were available. Fortunately, the UK is not a country where VAAs are very popular. Still, about 17% of control participants reported that they used a VAA over the course of the campaign[7], which could lead to downward bias. In Online Appendix §2.10, we re-estimate all models using two-stage least squares regressions where the first stage predicts actual VAA usage using treatment assignment as an instrument, and the second stage estimates the effects of VAA usage on our various outcome variables. While complier effects tend to be somewhat larger, the differences are marginal. Notably, the average effect on actual vote choice remains statistically insignificant, and we continue not to find evidence for an effect on electoral participation.

[7] This figure seems high for the UK, but members of online access panels are likely to have a high degree of Internet affinity.

We report several other robustness checks in Online Appendix §2.10. First, we drop speeders who rushed through the second panel wave, when treatment was applied. Second, we re-estimate all models with binary outcome variables using logistic regression. Third, we restrict the sample to participants who participated in all four panel waves. Finally, we show the effects of VAA usage on perceived policy proximity and party preferences when we do not interact treatment status with the ordinal ranking of parties in the VAA from closest to most-distant on policy issues, but with the underlying continuous (-100 to +100) proximity scores. The results are always similar.

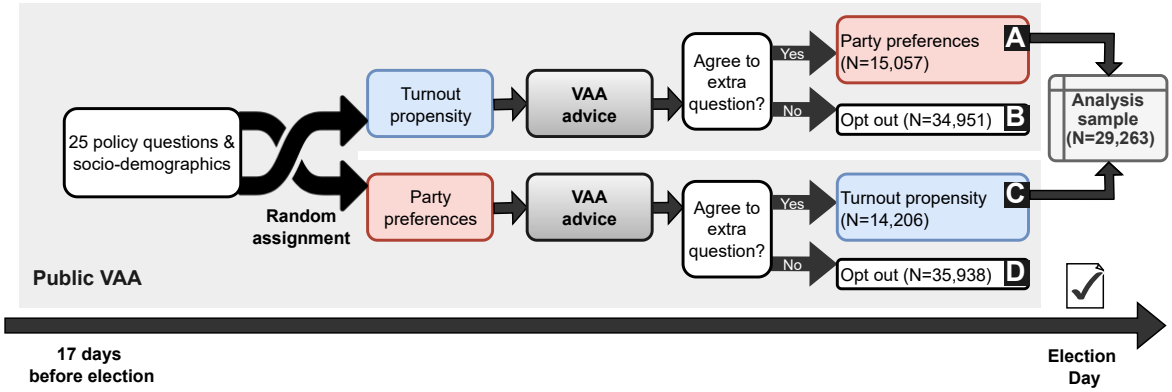
Study 2

An important remaining question is to what extent the effects of artificially induced VAA usage generalize to VAA usage in the real world. In study 2, we draw on data from a concurrent field experiment to shed light on the generalizability of the effects we measure in study 1.

Experimental Design

Study 2 leverages the circumstance that the VAA we administered in study 1 was simultaneously released to the general public. The public version was made freely available online 17 days before the 2024 UK General Election and promoted via print, broadcast, online, and social media, including paid advertisements on Facebook, Instagram, and X/Twitter. It was used around 125,000 times to generate a voting recommendation. After data cleaning (e.g., removal of repeated entries from the same individuals), we are left with a dataset comprising approximately 100,000 British voters (see Online Appendix §3.1 for details). However, for reasons we explain below, we only analyze a subset of approximately 29,000 users. Our VAA sample, including our smaller analysis sample, significantly over-represents highly educated voters and, to a lesser extent, younger voters (see Online Appendix §3.2). These biases are relatively typical of VAA user samples (Marschall and Schmidt 2008; Vassil 2011). While our study 2 sample is therefore clearly not representative of the British adult population, it broadly mirrors the population of VAA users.

Figure 8: Overview of study 2



The public release of *WhoGetsMyVote* and the version completed by study 1 participants coexisted as broadly identical websites on different domains. Relevant differences between the two versions included, first, that the public release contained several additional questions to gather information on users' socio-demographic profile. We leverage this information for the measurement of covariates (see Online Appendix §3.3). Second, the public release did not include the pop-up message with the survey completion code, but instead included an experimental manipulation that was built directly into the VAA. Specifically, we randomized the *time* when VAA users were asked about their party preferences and turnout propensity. One experimental group was asked about their party preferences *before* exposure to the VAA advice, and their turnout propensity *after* exposure to the VAA advice. This order was reversed for the second experimental group (see Figure 8).

Randomizing whether outcome questions are asked before or after exposure to VAA advice allows us to estimate the *immediate* causal effects of exposure to VAA advice. Analogously to study 1, we measure party preferences by asking users how likely it is that they will ever vote for different parties on a scale of 0 to 10.[8] Also similar to study 1, we measure electoral participation by asking VAA users to rate the likelihood with which they will participate in the 2024 UK General Election on a scale of 0 to 10. Since this version of the VAA was released to the general public, we kept the platform user-friendly and limited the number of questions. Hence, we did not also ask users about their binary vote or turnout intention. For the same reason, we did not include a question on perceived proximity on

[8] Users were asked to rate 5 to 6 parties, depending on the region. Unlike in study 1, we did not ask Scottish users to rate Alba.

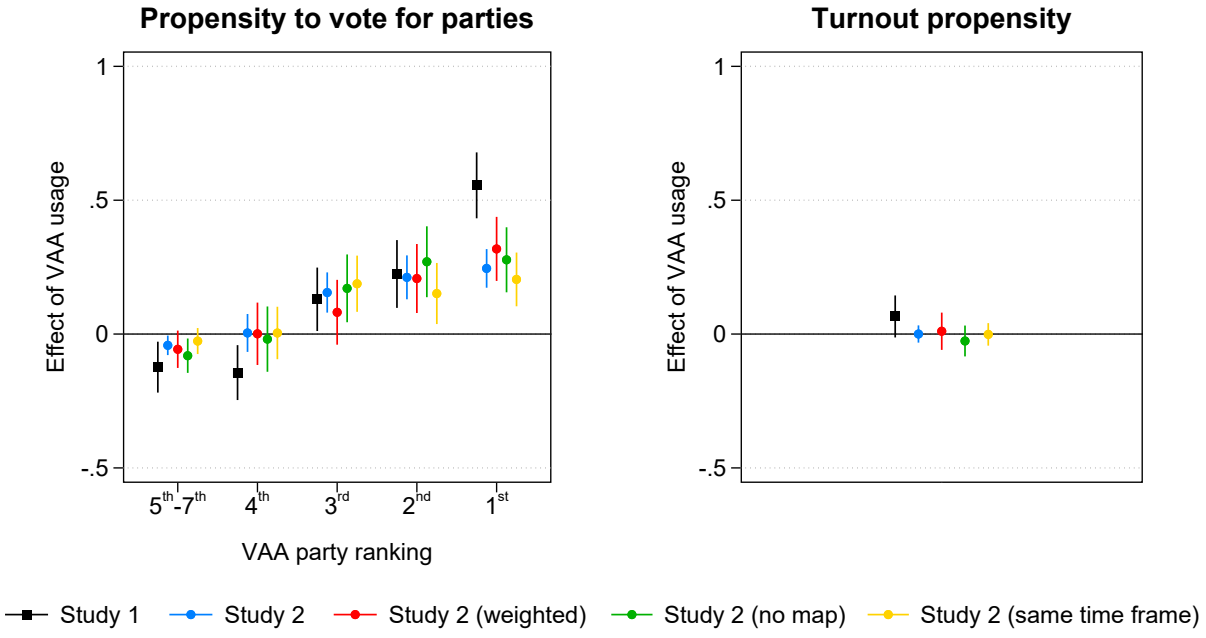
policy issues. While we are therefore unable to replicate the full set of study 1 results, we can contrast the immediate effects of artificially induced and real-world VAA usage on two important outcomes: the propensity to participate in elections and the propensity to vote for different parties. Online Appendix §3.4 contains the exact wordings of both outcome questions as well as summary statistics.

Finally, there is likely to be differential attrition between experimental groups (Garry et al. 2019; Germann, Mendez, and Gemenis 2023). We measure outcomes after exposure to the VAA advice via a pop-up window that appeared 30 seconds into the VAA advice screen. First, users were asked whether they would be willing to answer an additional question. If they said yes ("opted in"), users were then queried about their party preferences or turnout propensity, depending on their experimental group. Similar to previous timing studies, around a third of VAA users agreed to answer an additional question while the remaining two-thirds either declined or had already left the page. In contrast, users were much more likely to provide valid responses before exposure to the VAA advice. Hence, we restrict our analysis to the third of users who participated in the opt-in survey (i.e., we compare groups A and C in Figure 8). Users from these two groups had the same experience, except that one was asked about party preferences before exposure to the VAA advice and the other about turnout propensity. As a result, they were equally likely to participate in the opt-in survey (28% vs. 30%). Restricting the sample to users who participated in the opt-in survey removes potential bias due to differential attrition. Consistent with this, we find no differences between groups A and C in terms of both their socio-demographic profile and the VAA advice they received (see Online Appendix §3.5).

Results

We estimate the effects of exposure to VAA advice analogously to study 1, except that the treatment is now exposure to VAA advice in the real world and not artificially induced VAA usage. Similar to study 1, exposure to VAA advice leads real-world VAA users to adjust their party preferences in line with the VAA advice (see Figure 9). Also similar to study 1, we find no evidence for an effect on electoral participation. Overall, these results significantly strengthen in the generalizability of our study 1 results. That said, it is notable

Figure 9: Comparing the effects of artificially induced VAA usage with concurrent real-world effects



Note: The spikes represent 95% confidence intervals.

that the real-world effects on party preferences are somewhat smaller compared to the effects of artificially induced VAA usage. Most notably, while study 1 suggests that a VAA user's propensity to vote for the top-ranked party increases by about 0.6 points ($p < 0.001$), study 2 suggests an increase that is about half as large ($p < 0.001$).

Several factors are likely to account for the smaller effect sizes in study 1. First, one, albeit more minor, factor is likely that the study 2 sample underrepresents voters who are more strongly affected by VAA usage. In Figure 9, we include the results when the study 2 sample is weighted such that it matches the study 1 sample in terms of a range of potential drivers of treatment effect heterogeneity, including whether users already had a vote intention before VAA usage, as well as their age, education, and political interest. The effect sizes increase somewhat after weighting (see the red estimates), suggesting that the study 2 sample somewhat underrepresents voters who are likely to be strongly affected by VAAs. However, the difference to the unweighted results (blue estimates) is relatively marginal. Differences in sample composition are therefore unlikely to constitute the sole (or even the main) explanation.

Another contributing factor, again more minor, is likely that the treatment in study 2

had a slightly more compound nature. Both study 1 and study 2 participants were initially presented with a ranked bar chart. However, real-world VAA users could also access a second graphical display as part of the VAA results: an ideological map placing both the VAAs user and the different parties in a two-dimensional space. While the bar chart and the map tended to suggest similar conclusions, there can be minor differences since the map is based on a selection of the policy items. As we are estimating whether VAA usage increased alignment with the primary VAA result (i.e., the bar chart), this could lead to a small downward bias in study 2. Indeed, we find some evidence for this: as the green estimates show, the effect sizes are somewhat larger when the study 2 sample is restricted to the third of users who had only seen the bar chart and not the two-dimensional map by the time they filled in the outcome questions. However, the increases are relatively marginal, suggesting that other factor(s) are likely to play a bigger role.

A third explanation could be that the size of VAA effects varies over the course of an election campaign. However, we get broadly simila results when we limit the study 2 sample to individuals who completed the VAA during the time that treatment was applied in study 1 (see the gold estimates). This therefore leaves a fourth and final possibility: study 1 participants paid more attention to the VAA advice and/or had a better understanding of the results. While we are unable to directly test this notion, previous research suggests that Prolific panelists tend to have comparatively high attention and comprehension levels (Douglas, Ewell, and Brauer 2023; Peer et al. 2022). Unlike real-world VAA users, study 1 participants also received a financial reward, which could have further increased attention levels. This suggests that lower attention and comprehension likely represent a key reason why the effects in study 2 are smaller than in study 1. Nonetheless, and this is the most important result: we find meaningfully sized effects on party preferences in both studies. While VAA effects may therefore be larger among Prolific panelist, they are not limited to them.

Conclusion

The results of study 1 show that VAAs fulfill their primary aim: they help voters better understand where parties stand relative to their own policy positions. Furthermore, study 1 suggested that voters take the advice received seriously enough to revise their party preferences, though that effect is notably smaller compared to the effect on perceived proximity, most likely because party preferences reflect additional considerations. Both of these effects decrease over time, but they remain visible throughout the campaign and after the election. Finally, VAA usage influenced vote choice among undecided voters, suggesting that VAAs help some voters make choices that are more closely aligned with their policy preferences.

These results diverge from most earlier experimental literature, which tended to find limited or no causal effects of VAA usage. We argue that this is likely due to methodological limitations in prior work. Study 1 leveraged a new experimental design that ensures significantly higher treatment compliance and, even more importantly, allowed us to evaluate VAA effects on party evaluations conditional on the VAA advice received by participants. In contrast, most previous studies lacked access to the VAA advice and therefore estimated effects of VAA usage on vote switching, which is likely to obscure VAA effects. As we showed, VAA usage has clear and durable effects on party evaluations when the nature of the advice received is taken into account. That said, we found no evidence that VAA usage increases turnout. Hopes that VAAs could help reverse increasing electoral disengagement may thus be overstated.

To assess the ecological validity of our findings, we conducted a simultaneous field experiment using the public version of the same VAA. The effects observed in the field were somewhat smaller, but the overall pattern was consistent: VAA usage affected party preferences, but not turnout. Although we could not directly assess perceived policy proximity in the field, the observed shifts in preferences reflect greater understanding of parties' policy positions. Overall, our findings from study 2 suggest that VAA effects are not limited to artificially settings, but extend to real-world usage.

Taken together, our two studies point to the value of VAAs. Many voters have limited knowledge of parties' policy stances and biased views of their proximity to parties.

Our results suggest that VAAs can help rectify this. Critically, compared to many other interventions aimed at increasing political knowledge, VAAs are relatively inexpensive to produce, and, given their online nature, are easily scalable (Germann and Gemenis 2019). In a number of countries, VAAs are already used by hundreds of thousands or even millions of voters. Wider promotion of VAAs by researchers, governments, the media, or civil society organizations could help spread their benefits even further. By increasing voters' understanding of policy congruence with parties or candidates, VAAs ultimately have the potential to contribute to higher democratic quality.

There are several promising avenues for future research. First, the evidence we presented pertains to a single country: the United Kingdom. We argued that the UK represents an advantageous case to study since VAAs are currently not as popular in the UK, thus reducing contamination in (study 1's) control group. At the same time, though, the UK's majoritarian electoral system imposes significant constraints on voters. We presented tentative evidence suggesting that VAAs may have up to twice as large effects on vote choice in countries with more proportional electoral systems, where voters are freer to express their party preferences. Future comparative studies should explore the context-dependence of VAA effects in more detail.

Second, we showed that VAA effects endure, at least partially, over the course of an election campaign. However, our study cannot answer the question whether VAAs also have more long-term effects. Future studies could explore whether VAAs continue to affect how voters think about politics in the next election, and beyond. If they do, this could make VAAs a particularly valuable tool for school settings (Waldvogel, Oberle, and Leunig 2023).

Third, our results suggest that low attention levels may depress the effects of VAA usage. This points to the need for research on how VAAs can be designed to maximize users' attention levels and information retention. It could be particularly promising to explore the integration of Artificial Intelligence into VAAs, for example, as a way to help VAA users gain a better understanding of the policy issues they are asked about (Gemenis 2024) or to dynamically optimize and, thus, shorten VAA questionnaires (Bachmann, Sarasua, and Bernstein 2024; Sigfrid 2024).

Finally, our findings underline the importance of accurate VAA advice. Previous lit-

erature suggests that the quality of VAA advice depends on the selection and wording of policy questions (Gemenis 2013; Walgrave, Nuytemans, and Pepermans 2009), the coding of party (or candidate) positions (Gemenis and van Ham 2014), and how user and party (or candidate) preferences are aggregated and compared (Germann et al. 2015; Mendez 2012). If VAAs shape voters' understanding of politics and their electoral choices, then this creates a strong ethical obligation for VAA developers to follow best practices in VAA design.

References

Achen, Christopher H., and Larry M. Bartels. 2017. *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton, NJ: Princeton University Press.

Adams, James, Lawrence Ezrow, and Zeynep Somer-Topcu. 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.

Alvarez, R. Michael. 1998. *Information and Elections.* Ann Arbor, MI: University of Michigan Press.

Alvarez, R. Michael, Ines Levin, Peter Mair, and Alexander H. Trechsel. 2014. Party Preferences in the Digital Age: The Impact of Voting Advice Applications. *Party Politics* 20 (2): 227–236.

Andreadis, Ioannis, and Matthew Wall. 2014. The Impact of Voting Advice Applications on Vote Choice. In *Matching Voters with Parties and Candidates,* edited by Diego Garzia and Stefan Marschall, 105–128. Colchester: ECPR Press.

Bachmann, Fynn, Cristina Sarasua, and Abraham Bernstein. 2024. Fast and Adaptive Questionnaires for Voting Advice Applications. In *Machine Learning and Knowledge Discovery in Databases: Applied Data Science Track,* edited by Albert Bifet, Tomas Krilavičius, Ioanna Miliou, and Slawomir Nowaczyk, 365–380. Cham: Springer Nature Switzerland.

Benesch, Christine, Rino Heim, Mark Schelker, and Lukas Schmid. 2023. Do Voting Advice Applications Change Political Behavior? *Journal of Politics* 85 (2): 684–700.

Borelli, Gabriel. 2023. *Americans Differ over How Important It Is for Political Candidates They Support to Thare Their Personal Traits.* Pew Research Center, October 3.

Brown, John. 1958. Some Tests of the Decay Theory of Immediate Memory. *Quarterly Journal of Experimental Psychology* 10 (1): 12–21.

Clifford, Scott, Geoffrey Sheagley, and Spencer Piston. 2021. Increasing Precision without Altering Treatment Effects: Repeated Measures Designs in Survey Experiments. *American Political Science Review* 115 (3): 1048–1065.

Cohen, Geoffrey L. 2003. Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs. *Journal of Personality and Social Psychology* 85 (5): 808–822.

Dalager, Jon K. 1996. Voters, Issues, and Elections: Are the Candidates' Messages Getting Through? *Journal of Politics* 58 (2): 486–515.

De Graaf, Jochum. 2010. The Irresistible Rise of the StemWijzer. In *Voting Advice Applications in Europe: The State of the Art,* edited by Lorella Cedroni and Diego Garzia, 35–46. Naples: ScriptaWeb.

Douglas, Benjamin D., Patrick J. Ewell, and Markus Brauer. 2023. Data Quality in Online Human-Subjects Research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLoS ONE* 18 (3): e0279720.

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

Enyedi, Zsolt. 2016. The Influence of Voting Advice Applications on Preferences, Loyalties and Turnout: An Experimental Study. *Political Studies* 64 (4): 1000–1015.

Fawcett, Jonathan M., and Justin C. Hulbert. 2020. The Many Faces of Forgetting: Toward a Constructive View of Forgetting in Everyday Life. *Journal of Applied Research in Memory and Cognition* 9 (1): 1–18.

Fieldhouse, Ed, Jane Green, Geoffrey Evans, Jon Mellon, Chris Prosser, Roosmarijn de Geus, and Jack Bailey. 2022. *British Election Study, 2019: Post-Election Random Probability Survey.* UK Data Service.

Garry, John, James Tilley, Neil Matthews, Fernando Mendez, and Jonathan Wheatley. 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, Diego, Andrea de Angelis, and Joëlle Pianzola. 2014. The Impact of Voting Advice Applications on Electoral Participation. In *Matching Voters with Parties and Candidates,* edited by Diego Garzia and Stefan Marschall, 105–114. Colchester: ECPR Press.

Garzia, Diego, Alexander H. Trechsel, and Andrea de Angelis. 2017. Voting Advice Applications and Electoral Participation: A Multi-Method Study. *Political Communication* 34 (3): 424–443.

Gemenis, Kostas. 2013. Estimating Parties' Positions through Voting Advice Applications: Some Methodological Considerations. *Acta Politica* 48 (3): 268–295.

———. 2024. Artificial Intelligence and Voting Advice Applications. *Frontiers in Political Science* 6:1286893.

Gemenis, Kostas, and Martin Rosema. 2014. Voting Advice Applications and Electoral Turnout. *Electoral Studies* 36:281–289.

Gemenis, Kostas, and Carolien van Ham. 2014. Comparing Methods for Estimating Parties' Positions in Voting Advice Applications. In *Matching Voters to Parties and Candidates,* edited by Diego Garzia and Stefan Marschall, 33–47. Colchester: ECPR Press.

Germann, Micha, and Kostas Gemenis. 2019. Getting out the Vote with Voting Advice Applications. *Political Communication* 36 (1): 149–170.

Germann, Micha, Fernando Mendez, and Kostas Gemenis. 2023. Do Voting Advice Applications Affect Party Preferences? Evidence from Field Experiments in Five European Countries. *Political Communication* 40 (5): 596–614.

Germann, Micha, Fernando Mendez, Jonathan Wheatley, Constantinos Djouvas, Roula Nezi, and Matthew Wall. 2024. *Improving Issue Representation with Candidate-Level Voting Advice Applications.* Mimeo, University of Bath.

Germann, Micha, Fernando Mendez, Jonathan Wheatley, and Uwe Serdült. 2015. Spatial Maps in Voting Advice Applications: The Case for Dynamic Scale Validation. *Acta Politica* 50 (2): 214–238.

Grand, Peter, and Guido Tiemann. 2013. Projection Effects and Specification Bias in Spatial Models of European Parliament Elections. *European Union Politics* 14 (4): 497–521.

Green, Jane. 2007. When Voters and Parties Agree: Valence Issues and Party Competition. *Political Studies* 55 (3): 629–655.

Hanretty, Chris, Jonathan Mellon, and Patrick English. 2021. Members of Parliament are Minimally Accountable for Their Issue Stances (and They Know It). *American Political Science Review* 115 (4): 1275–1291.

Heinsohn, Till, Matthias Fatke, Jonas Israel, Stefan Marschall, and Martin Schultze. 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.

Joslyn, Mark R. 2003. The Determinants and Consequences of Recall Error About Gulf War Preferences. *American Journal of Political Science* 47 (3): 440–452.

Kamoen, Naomi, Bregje Holleman, André Krouwel, Jasper van de Pol, and Claes De Vreese. 2015. The Effect of Voting Advice Applications on Political Knowledge and Vote Choice. *Irish Political Studies* 30 (4): 595–618.

Keele, Luke, and Jason W. Morgan. 2016. How Strong Is Strong Enough? Strengthening Instruments through Matching and Weak Instrument Tests. *Annals of Applied Statistics* 10 (2): 1086–1106.

Kleinnijenhuis, Jan, Jasper van de Pol, Anita M.J. van Hoof, and André P.M. Krouwel. 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.

Kruikemeier, Sanne, Guda van Noort, Rens Vliegenthart, and Claes de Vreese. 2014. Unraveling the Effects of Active and Passive Forms of Political Internet Use: Does It Affect Citizens' Political Involvement? *New Media & Society* 16 (6): 903–920.

Lau, Richard R, and David P Redlawsk. 2006. *How Voters Decide: Information Processing in Election Campaigns.* Cambridge: Cambridge University Press.

Lau, Richard R., Parina Patel, Dalia F. Fahmy, and Robert R. Kaufman. 2014. Correct Voting Across Thirty-Three Democracies: A Preliminary Analysis. *British Journal of Political Science* 44 (2): 239–259.

Lewis-Beck, Michael S., and Mary Stegmaier. 2018. Economic Voting. In *The Oxford Handbook of Public Choice,* edited by Roger D. Congleton, Bernard Grofman, and Stefan Voigt, 247–265. Oxford: Oxford University Press.

Lodge, Milton, Marco R. Steenbergen, and Shawn Brau. 1995. The Responsive Voter: Campaign Information and the Dynamics of Candidate Evaluation. *American Political Science Review* 89 (2): 309–326.

Mahéo, Valérie-Anne. 2016. The Impact of Voting Advice Applications on Electoral Preferences: A Field Experiment in the 2014 Quebec Election. *Policy & Internet* 8 (4): 391–411.

———. 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, Vasilis, Fernando Mendez, and Constantinos Djouvas. 2020. *Mapping Voters' Engagement with Political Information Sites: The Case of Voting Advice Applications.* Paper presented at the 14$^{\text{th}}$ General Conference of the European Consortium of Political Research, August 2020.

Marschall, Stefan, and Christian Schmidt. 2008. Preaching to the Converted or Making a Difference? Mobilizing Effects of an Internet Application at the German General Election 2005. In *Non-Party Actors in Electoral Politics,* edited by David Farrell and Rüdiger Schmitt-Beck, 259–278. Baden-Baden: Nomos.

Marschall, Stefan, and M. Schultze. 2012. Voting Advice Applications and their Effect on Voter Turnout: The Case of the German Wahl-O-Mat. *International Journal of Electronic Governance* 5 (3–4): 349–366.

Mendez, Fernando. 2012. Matching Voters with Political Parties and Candidates: An Empirical Test of Four Algorithms. *International Journal of Electronic Governance* 5:264–278.

Merrill, Samuel III, Bernard Grofman, and James Adams. 2001. Assimilation and Contrast Effects in Voter Projections of Party Locations: Evidence from Norway, France, and the USA. *European Journal of Political Research* 40 (2): 199–223.

Miller, Beth. 2013. Failing to Recall: Examining the Effects of Trace Decay and Interference on Memory for Campaign Information. *Political Psychology* 34 (3): 289–305.

Munzert, Simon, Pablo Barberá, Andrew Guess, and JungHwan Yang. 2020. Do Online Voter Guides Empower Citizens? Evidence from a Field Experiment with Digital Trace Data. *Public Opinion Quarterly* 84 (3): 675–698.

Munzert, Simon, and Sebastian Ramirez-Ruiz. 2021. Meta-Analysis of the Effects of Voting Advice Applications. *Political Communication* 38 (6): 691–706.

Murray, Michael P. 2006. Avoiding Invalid Instruments and Coping with Weak Instruments. *Journal of Economic Perspectives* 20 (4): 111–132.

Peer, Eyal, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. 2022. Data Quality of Platforms and Panels for Online Behavioral Research. *Behavior Research Methods* 54 (4): 1643–1662.

Pianzola, Joëlle. 2014a. Selection Biases in Voting Advice Application Research. *Electoral Studies* 36:272–280.

———. 2014b. Swing Voting Due to Smartvote Use? Evidence from the 2011 Swiss Federal Elections. *Swiss Political Science Review* 20 (4): 651–677.

Pianzola, Joëlle, Alexander H. Trechsel, Kristjan Vassil, Guido Schwerdt, and R Michael Alvarez. 2019. The Impact of Personalized Information on Vote Entention: Evidence from a Randomized Field Experiment. *Journal of Politics* 81 (3): 833–847.

Reiljan, Andres, Frederico Ferreira da Silva, Lorenzo Cicchi, Diego Garzia, and Alexander H. Trechsel. 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* 31:105968.

Riera, Pedro. 2015. Tactical Voting. In *Oxford Handbook Topics in Politics.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199935307.013.55.

Rogers, Steven. 2017. Electoral Accountability for State Legislative Roll-Calls and Ideological Representation. *American Political Science Review* 111 (3): 555–571.

Schultze, Martin. 2014. Effects of Voting Advice Applications (VAAs) on Political Knowledge about Party Positions. *Policy & Internet* 6 (1): 46–68.

Sigfrid, Karl. 2024. IRT for Voting Advice Applications: A Multi-Dimensional Test that Is Adaptive and Interpretable. *Quality & Quantity* 58 (5): 4137–4156.

Tromborg, Mathias Wessel, and Andreas Albertsen. 2023. Candidates, Voters, and Voting Advice Applications. *European Political Science Review* 15 (4): 582–599.

Vassil, Kristjan. 2011. Voting Smarter? The Impact of Voting Advice Applications on Political Behavior. PhD diss., European University Institute.

Waldvogel, Thomas, Monika Oberle, and Johanna Leunig. 2023. What Do Pupils Learn from Voting Advice Applications in Civic Education Classes? Effects of a Digital Intervention Using Voting Advice Applications on Students' Political Dispositions. *Social Sciences* 12 (11): 621.

Walgrave, Stefaan, Michiel Nuytemans, and Koen Pepermans. 2009. Voting Advice Applications and the Effect of Statement Selection. *West European Politics* 32 (6): 1161–1180.

Walgrave, Stefaan, Peter van Aelst, and Michiel Nuytemans. 2008. 'Do the Vote Test': The Electoral Effects of a Popular Vote Advice Application at the 2004 Belgian Elections. *Acta Politica* 43 (1): 50–70.

Wall, Matthew, André Krouwel, and Thomas Vitiello. 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.

Wheatley, Jonathan, and Micha Germann. 2024. Do Candidates' Policy Positions Matter in Regional Elections? Evidence from the 2021 Elections to the Welsh Senedd. *Regional & Federal Studies,* 1–22.

Wixted, John T. 2004. The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology* 55:235–269.