**Measuring agreement: How to arrive at reliable measures of opinion congruence between voters and parties**

**Abstract**

The extent to which voters and parties agree on policies is an important way through which political scientists have empirically studied political representation. This opinion congruence is most often measured by comparing preferences on a number of policy statements. While the selection of policy statements has not escaped scholarly attention, its impact on the reliability of congruence scores, i.e. the degree to which similar levels of opinion congruence are found when different samples of policy statements are used, has been less investigated. This article looks at which factors of statements samples and voters affect the reliability of congruence measures. It does so by simulating over 5 million opinion congruence scores on the basis of a dataset containing 134 voter and party policy preferences. It finds that both the number of statements and their topic diversity positively affect the reliability of congruence estimates. In addition, the congruence estimates of politically less sophisticated voters are more reliable but only when many left-right policy statements are included in the statement selection. Finally, explorative analyses suggest that increasing topic diversity also increases the validity of congruence measures.

**Introduction**

Normative democratic theory emphasizes the proper representation of voters’ policy preferences by representatives. High levels of policy agreement between voters and elites, in what has been referred to as ‘opinion congruence’, is one of the most important indicators of democratic health (Dahlberg, Linde, & Holmberg, 2015; Diamond & Morlino, 2005), and a large body of research has been devoted to this topic (Walczak and Brug 2013; Walgrave and Lefevere 2013). The logic behind the study of opinion congruence is simple: when citizens and representatives have similar policy positions, there is an increased likelihood of these positions becoming actual policies (Dalton 1988; Thomassen 1994).

The measurement of opinion congruence has gone through several developments since Stokes and Miller’s (1963) seminal article about the correlation between the preferences of Congressional representatives and their constituents in the US. A first development happened when correlational measures were replaced by estimates of distance between voter and elite positions (Achen 1977). In addition, scholars have reconsidered how to measure policy positions. For a long time, agreement between voters and elites has been measured by examining congruence on aggregated measures of policy preferences such as the left-right scale, in what has been labelled ideological congruence (Andeweg 2011) or the ‘policy mood’ (Stimson, Mackuen, and Erikson 1995). By now, however, it has been well established that such global measures are insufficient to capture the policy preferences of voters, and by extension their opinion congruence with parties (Freire and Belchior 2013; Todosijević 2004). In response, most recent works on opinion congruence rely on comparisons between voters’ and elites’ positions on multiple policy items (Walgrave and Lefevere 2013; Lefkofridi, Wagner, and Willmann 2014; Lupu and Warner 2017; Giger, Rosset, and Bernauer 2012; Holmberg 2000). If it is becoming increasingly clear that global assessments of the opinion congruence between voters and political elites need to be based on a list of specific policy items, the next question then is what such a list must look like. What should a list of specific policy items consist of, if it aims to replace and outperform the left-right scale?

Here too, however, have scholars started to examine the validity of the opinion congruence measures, ranging from the formulation of the statements (see Camp, Lefevere, and Walgrave 2014), to the correlation between policy opinion congruence and ideological congruence (Freire and Belchior 2013), and how characteristics of the statements selected to measure policy opinion congruence affects the outcome (Lefevere and Walgrave 2014, 2015). As such, inquiries into the methodological aspects of opinion congruence has focused primarily on their validity, and less on their reliability, i.e. the degree to which similar levels of opinion congruence are found when different samples of policy statements are used.

In this paper, we want to fill this gap in the literature by exploring the factors that can make measures of the overall opinion congruence[[1]](#footnote-1) between voters and parties consistent. In examining factors that influence opinion congruence variability (OCV), we distinguish between characteristics of statements and voters. Specifically, we want to examine the relevance of five factors: 1) the number of policy statements, 2) the topic diversity of the policy statements, 3) whether statements can be classified in the left-right spectrum, 4) personal issue salience, and 5) voters’ political interest.

**The reliability of opinion congruence measures**

In our exploration of factors that reduce inconsistencies in opinion congruence measures, we begin with the most obvious two factors: the number of policy statements used to calculated opinion congruence, and the range of issue topics covered. Regarding the first, the law of large numbers argues that when an experiment is performed a large number of times, the average result will approximate the actual congruence on issues that currently make up public debates.[[2]](#footnote-2) Similarly, we can apply this line of reasoning to the measurement of opinion congruence. Each policy statement that is about a topical issue is an experiment of whether voters and parties agree or not. When adding more and more policy statements, the underlying pattern of congruence between their policy views should reveal itself. In other words, if we were to randomly draw more and more policy statements from the universe of debated policy issues and use those statements to calculate congruence, increasing the number of policy statements drawn brings the observed opinion congruence score closer to the true level of opinion congruence given contemporary public debates. Therefore, the larger the number of policy statements, the more likely congruence scores will be consistent with one another, even when calculated on the basis of a completely different set of policy statements.

However, adding more policy statements can still result in an unreliable estimate of opinion congruence if the diversity of topics covered remains limited. For instance, when voters and parties agree environmental policies, but disagree starkly on matters related to immigration. If the selection of policy statements ignores the latter topic, we arrive at an overly optimistic estimate of opinion congruence, an estimate that will differ from one which does take immigration into account. In this regard, Thomassen (2012) argued that limiting the topic scope could result in ‘blind corners’ in congruence research. In other words, opinion congruence measures are expected to be more robust and consistent when the policy statements on which the measures are based are spread across a wider variety of issue topics. The greater the diversity of topics included, the greater the nuance with which we measure whether voters and parties share the same views on policy.

Voters differ, however, in how much they know about policy issues and topics (Fowler and Margolis 2014), and this knowledge (or lack thereof) can affect estimations of opinion congruence. Voters who are less informed about a topic are also less likely to have stable and by extension internally consistent preferences, in contrast to elites and parties who usually have a very consistent belief system (see Converse 1964; Todosijević 2004). Consequently, calculating opinion congruence between voters and parties will result in varying outcomes. Two important factors that affect how much voters know about a topic is personal topic saliency and a general interest in politics. While the importance of saliency in opinion congruence has been stressed decades ago (Kuklinski and McCrone 1980), only recently have scholars begun to integrate it in earnest (see for instance Giger and Lefkofridi 2014; Walgrave and Lefevere 2013). Saliency increases the motivation to search for policy-relevant information (Ciuk and Yost 2016; Krosnick 1988, 1990), which in turn increases in the internal consistency of voters’ attitudes (Judd and Downing 1990; Lavine, Thomsen, and Gonzales 1997). This attitudinal stability is expected to in turn have a stabilizing effect on estimations of opinion congruence between voters and parties.

For instance, if voters care little about climate change, their positions on policy statements related to climate change are more likely to resemble what Converse (1964) referred to as ‘nonattitudes’, attitudes given with only a vague understanding of what the question is about (Sturgis and Smith 2010). The resulting inconsistency in attitude direction is more likely to result in fluctuating congruence scores, sometimes being high or low levels depending on the policy statements sampled. In contrast, on salient issues attitudinal direction and by extension the estimations of party-voter opinion congruence are more likely to be stable, regardless of the specific policy statements used to calculate opinion congruence. As such, opinion congruence measures that integrate voters’ personal issue salience (see Lavine et al. 1996) into their calculations are predicted to be more stable.

In addition to differences in personal saliency between different topics, voters differ in general in their interest in political issues. A consistent finding in political science has been the large gap in political knowledge and interest between various segments of the voting public (Campbell et al. 1960; Kinder 2006; Zaller 1992). The arguments regarding the relation between issue salience and attitudinal stability apply here as well, with the difference that congruence measures are expected to be more stable for politically interested voters than they are for less interested voters, ceteris paribus. This is important, as a growing body of research has examined the differences or inequality in opinion congruence between various groups (Aaldering 2017; Schakel and Hakhverdian 2018a). Oftentimes, the groups distinguished also differ in terms of their interest in and knowledge about politics. The less politically sophisticated, in an effort to compensate for lack of issue-specific knowledge, are expected to be more reliant on heuristics. In this regard, the left-right ideological dimension is still the most important heuristic in politics and it remains the pivotal structuring schemata in many polities (Knutsen 1995). As such, voters’ positions on policy statements related to the central left-right divide should be more internally consistent (Milburn 1987), resulting in estimations of opinion congruence that are more reliable.

Based on the previous discussion, we can formulate the following hypotheses: H1) the reliability of opinion congruence estimates is expected to increase as the number of policy statements used increases. H2) As the diversity of topics or issue domains increases, so will the reliability of opinion congruence estimates. H3) Estimates of opinion congruence that take into account voters’ personal issue salience are expected to be more reliable than estimates that do not. H4a) Opinion congruence estimates of politically sophisticated voters are more reliable than estimates of less politically sophisticated voters, H4b) but this difference decreases as more policy statements are related to the central left-right divide.

**Data and Method**

We use two datasets to test our hypotheses. The first is an online voter survey of 1,053 voters in Belgium’s largest region, Flanders. Voters were quota sampled, and in all analyses survey weights were used to ensure its representativeness of the Flemish population in terms of gender, age, and education[[3]](#footnote-3) (Ansolabehere and Rivers 2013), and were asked to give their policy preferences on 134 policy statements, spread across 12 policy topics (see Table A2 Appendix). These statements aimed to capture relevant topics in the public debates in Flanders in 2014.[[4]](#footnote-4) In order to avoid respondent fatigue due to the large number of policy statements, the survey was split into two waves. Due to this strategy, the average length of one survey wave was only 15 minutes. The response rate for the entire survey (both waves) was 17% (12,421 voters were contacted).[[5]](#footnote-5)

Concurrently, party leaderships were asked to give their party’s positions on those same 134 policy statements, in the context of an online voting aid application (VAA). Both voters and leaderships could either (0) disagree or (1) agree with the policy statement. Surveying party leaderships is regarded as a fruitful approach in Belgium as they rather than candidates/MP’s determine parties’ policy positions (Deschouwer 2012). The leadership survey included six parties: *Groen* (greens)*, Spa* (social democrats)*, CD&V* (christen democrats)*, Open VLD* (liberals)*, NVA* (Flemish regionalists)*,* and *Vlaams Belang* (extreme right)*.*This approach is not without its pitfalls. Previous research on party responses in a VAA showed a tendency towards centrist answers (Gemenis and Ham 2014; Wagner and Ruusuvirta 2011). However, we maintain that the leadership answers can be considered valid estimations of the party’s positions because they were critically examined by a research team of political scientists, resulting in several changes in position, and were widely discussed in the media.

Using the responses of the party leadership raises the question of whether the positions of the party leaderships always matched those of their rank and file MPs. Arguably, one might expect a high level of opinion congruence between party leaderships and party MPs: Candidates are unlikely to join a party with which they starkly disagree, and parties are unlikely to allow a candidate to represent them who does not endorse the party leadership’s positions. In addition, even in the case of disagreements, there are still important reasons to assume that MPs will vote in line with the party leadership. These include anticipated sanctions and adherence to the norm dictating that MPs should express loyalty to the party leadership (Andeweg & Thomassen, 2011, Van Vonno et al, 2013). In sum, MPs and the party leadership are highly likely to agree on the vast majority of issues, but even when they do not, the latter’s position is the one that matters. The near-total party cohesion during votes in parliament has exemplified this in the Belgian case (Depauw 2003).

With this dataset, we simulate estimates of opinion congruence. These consist of five steps. In the first, a sample of size *s* policy statements is drawn from the list of 134 statements (S), and opinion congruence is calculated for each voter-party combination[[6]](#footnote-6). This entails calculating the percentage of policy statements on which a voter and a party agree by dividing the number of policy statements on which a voter and a party agree by the number of statements sampled (*s*). In the second step, another sample of size *s* policy statements is drawn from the remaining statements (i.e. the list of 134 minus the statements that were drawn in the first step, or S-s). For this second sample of statements, we also calculate opinion congruence for each voter-party combination. In a second version of the calculations, voters’ topic salience was integrated in the formula. In the survey, respondents were asked to indicate how important each of the 12 issue topics were to them on an 11-point scale. The calculation of the unweighted and weighted party-voter congruence is shown in Formulae 1 and 2.

To assess the variability of these scores, we take the absolute difference between the opinion congruence scores of the first sample and those of the second sample in a third step (Formula 3). The larger the difference, the more dissimilar and variable the two measurements of congruence are. In contrast, smaller differences indicate measurements that are more reliable. In order to get an accurate picture of OCV at *s* policy statements, we repeat this process 50 times in a fourth step. The sampling algorithm was designed in such a way that there were no duplicate comparisons. Finally, steps one through four are repeated in step five, each time increasing the size of *s* (the number of statements drawn in step one and two) until *s* = 66.

Formula 1: Unweighted Party-voter opinion congruence between party A and voter J

$$Unweighted Opinion congruence=\frac{\sum\_{i=1}^{s}1-\left|policy position\_{i} of party A-policy position\_{i} of voter J\right|}{s}$$

Formula 2: Weighted party-voter opinion congruence between party A and voter J

$$Weighted opinion congruence=\frac{\left(\sum\_{i=1}^{s}\frac{\left(1-\left|policy position\_{i} of party A-policy position\_{i} of voter J\right|\right)\*saliency\_{i}}{\sum\_{i=1}^{s}saliency\_{i}}\right)}{s}$$

Formula 3:

$$Opinion congruence variability =\left|Opinion congruence sample 1-Opinion congruence sample 2\right|$$

These simulations give us a rich dataset in which we have, for every respondent, 50 pairs of statement samples, on the basis of which we are able to calculate opinion congruence variability for every of the six parties, for increasing sizes of the statement samples (1 to 66). In this dataset, respondents and parties are cross-nested within each pair of samples and repeated across many samples in a stacked dataset of roughly 10 million observations (after removing missing values), given that the opinion congruence scores are once unweighted and once weighted with their salience to voters. Figure 1 gives an overview of the simulation process, and an example of the dataset is shown in Table A4 the Appendix. Because of the nested nature of the data, we adopt a multi-level model with a random-intercept at the voter level (Gelman and Hill 2006). With only six political parties, it makes little sense to add an additional random intercept on the party-level due to limited variance between parties. To control for party differences, all models include party dummies.

[insert Figure 1 here]

The first independent variable is the *number of policy statements* drawn in step one and two of the calculation process, which varies between 2 and 66. The second independent variable is the *topic diversity* of the statement samples. We measure this diversity through the Herfindahl–Hirschman Index (HHI) (Djolov 2013). This index indicates how equally represented the twelve issue topics are in each of the two policy statement samples. We take the reverse of the HHI so that high values indicate high level of topical diversity and vice versa. Topic diversity is operationalized as the smallest HHI of the two policy statement samples. Thus, topic diversity has high values when both statement samples are equally spread across many topic issues and low values when both samples or one of the samples is concentrated and touch upon only a few issue topics.

In addition, it is reasonable to assume that the marginal effect of the number of policy statement and the topic diversity decreases as the overall number and diversity grows larger, in what economists have labelled as the law of diminishing returns (Radelet, Clemens, and Bhavnani 2005). Therefore, we take the cube root of the number of policy statements and topic diversity. Additional analyses (reported in Table A5 in the Appendix), indicate that the cube-root transformation of the first two independent variables indeed provided a better fit with the data. We take the cube root instead of the natural log due to the inability of the latter transformation to handle zeros. In addition, we explored whether the relation for the other variables were also non-linear, but this was not the case.

The third independent variable, *weighted*, is a dummy variable that indicates whether the ORC is derived from opinion congruence scores weighted by voters’ personal salience (1) or not (0). The fourth independent variable indicates voters’ *political interest* and is based on a self-reported response on an 11-point scale. The fifth and final independent variable, *left-right statements*, measures the extent to which two samples of policy statements contained statements that relate the left-right dimension. This dimension consists of a socio-economic and cultural dimension (Hooghe, Marks, and Wilson 2002; Inglehart 1990). For our purposes, to which dimension a policy statement belongs is of less importance. As long as it belongs to one, voters are given a cue that anchors their position. No dimension could be assigned in more than 20 percent of the cases[[7]](#footnote-7). Therefore, for both samples of policy statements, we calculated the percentage of left-right statements. The percentage in the first sample was then multiplied with the percentage in the second sample, resulting in the variable Left-right statements, which has high values when both statement samples contain many left-right statements and low values when both samples have few left-right statements. In our analyses, we also control for education level gender, age, and income (measured as their income decile).

While the data spans a large number of policy domains, some policy domains are inevitably more present that others. As such, it is unavoidable that larger samples of policy statements will be more similar to each other in terms of the policy domains they cover. In order to make sure that it is size of the statement samples and the topic diversity that drives reductions in variability, and not the increased substantive similarity of sample pairs, we control for *policy domain difference*. This variable indicates the degree to which two samples of policy statements differ in the distribution of policy domains. For instance, if both statement samples consist of 10 statements that both cover education and the environment, but sample 1 has 5 statements for each and sample 2 has 8 education statements and 2 environment statements, the policy domain difference = | 5 – 8| + | 5 – 2| = 3 + 3 = 6. In other words, the larger the values for policy domain difference, the larger the differences in the distribution of policy domains between the two samples of policy domains that are being compared. How large this difference can get is naturally dependent on the size of the statement samples. To account for this, the policy domain difference was divided by the maximum possible value[[8]](#footnote-8).

**Results**

The hypotheses are tested in Table 1. Model 1 indicates highly significant effects of the number of policy statements, topic diversity, and personal salience. In other words, increasing the number of statements and topic diversity makes opinion congruence measures more reliable, but the strength of their effects diminishes. However, given the large sample size, statistical significance is not a useful metric by which to assess the support hypotheses received given the data. While a non-significant finding is evidence for a lack of support, statistical significance must be supplanted by a substantial effect size before we can conclude that an assertion is confirmed. To inspect the effect sizes, we plot the marginal effects of Model 1 in Figures 2-4.

In Figure 2, we see that the estimated difference between opinion congruence measures decreases rapidly as the number of policy statements increases. However, the marginal decreases grow smaller with each additional statement added to the sample. To further disentangle the curve-linear relation between the number of policy statements and congruence reliability, we have examined the cut-off points at which the association between sample size and variability changes. We divide the curved relation in Figure 2 into three parts, using iterative fitting with segmented regression. The best fitting model was one in which the relation between statement sample size and OCV was modelled to alter after the number of policy statements increases beyond 6 and 25. The coefficients of the model can be found in the Appendix in Table A6, but the results are visualized in Figure 3. In the first part of the curve (statement sample size 1 through 5), OCV declines rapidly, with opinion congruence variability declining with 1 point every 0.25 statements. In the second part of the curve (statement sample size 6 through 19), OCV declines with 1 point every 1.32 statements. In the last part (statement sample size 20 through 66), further substantial improvements in the reliability of congruence measures require a high number of addition policy statements, with OCV declining 1 point if 8 policy statement are added.

[Insert Table 1 here]

[insert Figure 2 here]

[insert Figure 3 here]

Looking at topic diversity in Figure 4, we see that when topic diversity is low, it is unlikely that two measures of opinion congruence will be anywhere near each other. The effect of topic diversity is even more clearly non-linear, with decreasing marginal improvement to congruence reliability. While Model 1 indicates a significant difference in OCV when opinion congruence is measured by personal salience, this effect is in opposite direction as predicted in hypothesis 3. However, plotting the distribution of both weighted and unweighted OCV shows that the difference is very small (Figure 5). In light of this, we would be hard-pressed to conclude that these results support hypothesis 3, in contrast to strong evidence in favor of hypotheses 1 and 2[[9]](#footnote-9).

[insert Figure 4 here]

[insert Figure 5 here]

In Model 1, the coefficient of political interest is not significant, leading us to reject hypothesis 4a. However, Model 2 reveals an interaction effect between political interest and the number of left-right statements in the samples, shown in Figure 4. The main effect of political interest is negative, indicating that, when opinion congruence is calculated on the basis of not a single left-right statement, OCV is negatively dependent on political interest. The direction of the interaction effect is positive, meaning that increasing the number of left-right statements in the policy statement samples lessens the negative effect of political interest. In other words, when the vast majority of statements can be linked to the left-right ideological spectrum, the distinction between voters on the basis of interest in politics with regards to OCV begins to fade, though it never fully disappears. While the moderating effect in Figure 4 is smaller than the impact of the number of statements and topic diversity, the effect remains substantial. These analyses also reveal the number of left-right statements as another factor that affects OCV. From the direct effect in Model 1, we find that opinion congruence measures that are strongly embedded in the left-right scale are more reliable than those that are not, though the difference is not large.

[insert Figure 6 here]

**Do sample size, diversity, and salience increase opinion congruence validity?**

Our first three hypotheses predicted that increasing the number of policy statements, the diversity of the topics covered by the statements, and in the weighting of statements by their importance would increase the reliability of the opinion congruence estimates. In addition to reliability, their inclusion in the measurement of opinion agreement arguably makes the estimates more encompassing and nuanced. In other words, it is possible that sample size, topic diversity, and personal salience also improve the validity of opinion congruence. While a full validity examination is beyond the scope of the present study, we do present here an explorative analysis. Specifically, we examine whether statement sample size, topic diversity, and personal salience increase the likelihood that one will find patterns of congruence that make sense.

While we are aware that this reverses the scientific process, whereby a hypothesis is used to examine data instead of the other way around, we believe at least one hypothesis could lend itself for this purpose. The finding that higher educated voters are more likely than lower educated voters to vote for the party they agree most with, in what has often been referred to as ‘correct voting’, has received widespread support in the literature (Aaldering 2017; Lau et al. 2014; Lesschaeve 2017a; Walgrave and Lesschaeve 2018). As such, if the number of policy statements, topic diversity, and personal salience not only improve the reliability of congruence estimates, but also their validity, then they should increase the difference in correct voting between lower and higher educated. We test these assertions in Table 2.[[10]](#footnote-10) The dependent variable in these analyses is *correct voting*, which is the agreement between a voter and his/her vote intention for the 2014 elections, relative to the maximum degree of opinion congruence a voter could achieve, represented by the opinion congruence score of the party demonstrating the greatest agreement with the positions of a particular voter.

[Insert Table 2 here]

Model 1, which reports the direct effects of all variables on correct voting, confirms the findings of previous studies that higher educated voters are indeed more likely to vote for the party they agree with most. In addition, the model suggests that a greater numbers of policy statements, and a greater diversity in the topics covered substantially increase the estimates of correct voting. Personal salience lowers estimates of correct voting, but the effect is not sizeable. Model 2, which interacts education level with sample size, topic diversity, and topic salience, shows that only topic diversity significantly increases the correct voting gap between the lower and higher educated. In addition to a significant moderation, the interaction effect is also substantial. Figure 7 plots the predicted values for correct voting for all values of topic diversity, for lower, middle, and higher educated voters. It is clear that at low levels of topic diversity, correct voting hardly differs between the three groups, but that these differences grow substantially when the sample of statements covers a wider range of topics. This strongly suggests that, in addition to increasing the reliability of the congruence estimates, topics diversity increases their validity, though the same cannot be said about the number of policy statements and personal salience.

[insert Figure 7 here]

**Conclusion**

The goal of this paper was to get insights into how the opinion congruence between voters and parties is measured and contribute to its further development. Complementing existing studies on the validity of opinion congruence measures, this study examines which factors affect their reliability. Our results show that the number, the diversity, and the relation of policy statements to the left-right divide all substantially affect how reliable congruence measures are. On the voter level, we found differences when distinguishing on the basis of political interest, but only when few left-right policy statements were included. This last result is of particular importance to the strand in the literature that investigates diverging patterns of opinion congruence between groups of voters (Giger, Rosset, and Bernauer 2012; Rosset and Stecker 2019). In short, our study shows that proper measures of opinion congruence should be based on numerous policy statements, that are spread across a wide range of policy domains, and that connect to the main political cleavages in the polity (Lefevere and Walgrave 2015).

When it comes to making concrete recommendations to future policy statement-based research on opinion congruence, our analyses show that after roughly 20 statements, the number of additional policy statements needed to make further meaningful improvements in opinion congruence reliability increases exponentially. Besides that, scholars should aim to make their sample of policy statements cover a wide range of policy domains, and they should be cognizant of the possibility that the quality of their measures differs between groups of respondents. Integrating salience in opinion congruence, however, does not meaningfully improve the reliability and validity of the measures.

Though we think the evidence in this paper is compelling, this study is not without its limitations. Other, more aggregate conceptualizations of opinion congruence (Golder and Stramski 2010), might produce results different from the one’s reported here. In addition, our study only looks at one country, and thus only to one party system. Would the same difference between parties and voters be found in other party systems? We are not sure. Specifically, in Belgium politics are structured by a strong linguistic divide, in addition to the classic left-right cleavage (Deschouwer 2012; Freire 2015). It is possible that in countries with fewer cleavages, reliable congruence scores could be obtained with a lower number of policy statements. In addition, future research is needed to corroborate the findings presented here with other sources of party positions, such as manifesto data or candidate surveys.

In sum, we do not claim that our results are just transposable to other political and party systems. However, the results shown in this study offer a useful way to start thinking more systematically about how to measure opinion congruence. Empirical science depends on reliable measures and refining those measures is an integral part of the scientific process. We believe to have presented an approach that can lead to improved estimates of the agreement between voters and political parties, contributing to our understanding of opinion congruence and democratic representation.

**Cited literature**

Aaldering, Loes. 2017. “Political Representation and Educational Attainment: Evidence from the Netherlands (1994–2010).” *Political Studies* 65(1\_suppl): 4–23.

Achen, Christopher H. 1977. “Measuring Representation: Perils of the Correlation Coefficient.” *American Journal of Political Science* 21(4): 805.

Andeweg, Rudy. 2011. “Approaching Perfect Policy Congruence: Measurement, Development, and Relevance for Political Representation.” In *How Democracy Works: Political Representation and Policy Congruence in Modern Societies*, eds. Martin Rosema, Bas Denters, and Kees Aarts. Amsterdam: Amsterdam University Press, 39–52.

Andeweg, Rudy, and Jacques Thomassen. 2011. “Pathways to Party Unity: Sanctions, Loyalty, Homogeneity and Division of Labour in the Dutch Parliament.” *Party Politics* 17(5): 655–72.

Ansolabehere, Stephen, and Douglas Rivers. 2013. “Cooperative Survey Research.” *Annual Review of Political Science* 16(1): 307–29.

Camp, Kirsten Van, Jonas Lefevere, and Stefaan Walgrave. 2014. “The Content and Formulation of Statements in Voting Applications: A Comparative Analysis of 26 VAAs.” In *Matching Voters with Parties and Candidates. Voting Advice Applications in a Comparative Perspective*, eds. Diego Garzia and Stefan Marschall. Colchester: ECPR Press, 11–32.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. Unabridged ed. Chicago [Ill.]: University of Chicago Press.

Ciuk, David J., and Berwood A. Yost. 2016. “The Effects of Issue Salience, Elite Influence, and Policy Content on Public Opinion.” *Political Communication* 33(2): 328–45.

Converse, Philip E. 1964. “The Nature of Belief Systems in Mass Publics.” *Critical Review* 18(1–3): 1–74.

Dalton, Russell. 1988. *Citizen Politics in Western Democracies: Public Opinion and Political Parties in the United States, Great Britain, WestGermany, and France*. Chatham House Publishers.

Depauw, Sam. 2003. “Government Party Discipline in Parliamentary Democracies: The Cases of Belgium, France, and the United Kingdom in the 1990s.” *The Journal of Legislative Studies* 9(4): 130–46.

Deschouwer, Kris. 2012. *The Politics of Belgium: Governing a Divided Society*. Macmillan International Higher Education.

Djolov, George. 2013. “The Herfindahl-Hirschman Index as a Decision Guide to Business Concentration: A Statistical Exploration.” *Journal of Economic and Social Measurement* 38(3): 201–27.

Fowler, Anthony, and Michele Margolis. 2014. “The Political Consequences of Uninformed Voters.” *Electoral Studies* 34: 100–110.

Freire, André. 2015. “Left–Right Ideology as a Dimension of Identification and of Competition\*.” *Journal of Political Ideologies* 20(1): 43–68.

Freire, André, and Ana Belchior. 2013. “Ideological Representation in Portugal: MPs’–Electors’ Linkages in Terms of Left–Right Placement and Substantive Meaning.” *The Journal of Legislative Studies* 19(1): 1–21.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gemenis, Kostas, and Carolien van Ham. 2014. “Comparing Methods for Estimating Parties’ Positions in Voting Advice Applications.” In *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective*, eds. Diego Garzia and Stefan Marschall. Colchester: ECPR Press, 33–48.

Giger, Nathalie, and Zoe Lefkofridi. 2014. “Salience-Based Congruence Between Parties & Their Voters: The Swiss Case.” *Swiss Political Science Review* 20(2): 287–304.

Giger, Nathalie, Jan Rosset, and Julian Bernauer. 2012. “The Poor Political Representation of The Poor in a Comparative Perspective.” *Representation* 48(1): 47–61.

Golder, Matt, and Jacek Stramski. 2010. “Ideological Congruence and Electoral Institutions.” *American Journal of Political Science* 54(1): 90–106.

Holmberg, Soren. 2000. “ISSUE AGREEMENT MATTERS.” In *Beyond Westminster and Congress: The Nordic Experience*, Ohio State University Press, 155.

Hooghe, Liesbet, Gary Marks, and Carole J. Wilson. 2002. “Does Left/Right Structure Party Positions on European Integration?” *Comparative Political Studies* 35(8): 965–89.

Inglehart, Ronald. 1990. *Culture Shift in Advanced Industrial Society*. Princeton, N.J: Princeton University Press.

Judd, Charles M., and James W. Downing. 1990. “Political Expertise and the Development of Attitude Consistency.” *Social Cognition* 8(1): 104–24.

Kinder, Donald R. 2006. “Belief Systems Today.” *Critical Review* 18(1–3): 197–216.

Knutsen, Oddbjorn. 1995. “Value Orientations, Political Conflicts and Left-Right Identification: A Comparative Study.” *European Journal of Political Research* 28(1): 63–93.

Krosnick, Jon A. 1988. “The Role of Attitude Importance in Social Evaluation: A Study of Policy Preferences, Presidential Candidate Evaluations, and Voting Behavior.” *Journal of Personality and Social Psychology* 55(2): 196–210.

———. 1990. “Government Policy and Citizen Passion: A Study of Issue Publics in Contemporary America.” *Political Behavior* 12(1): 59–92.

Kuklinski, James H., and Donald J. McCrone. 1980. “Policy Salience and the Causal Structure of Representation.” *American Politics Quarterly* 8(2): 139–64.

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(02): 239–59.

Lavine, Howard, John L. Sullivan, Eugene Borgida, and Cynthia J. Thomsen. 1996. “The Relationship of National and Personal Issue Salience to Attitude Accessibility on Foreign and Domestic Policy Issues.” *Political Psychology* 17(2): 293.

Lavine, Howard, Cynthia J. Thomsen, and Marti Hope Gonzales. 1997. “The Development of Interattitudinal Consistency: The Shared-Consequences Model.” *Journal of Personality and Social Psychology* 72(4): 735–49.

Lefevere, Jonas, and Stefaan Walgrave. 2014. “A Perfect Match? The Impact of Statement Selection on Voting Advice Applications’ Ability to Match Voters and Parties.” *Electoral Studies* 36: 252–62.

———. 2015. “L’importance du choix des affirmations dans les SAV : résultats de deux simulations.” *Revue internationale de politique comparée* 23(2): 175.

Lefkofridi, Zoe, Markus Wagner, and Johanna E. Willmann. 2014. “Left-Authoritarians and Policy Representation in Western Europe: Electoral Choice across Ideological Dimensions.” *West European Politics* 37(1): 65–90.

Lesschaeve, Christophe. 2017a. “Inequality in Party-Voter Opinion Congruence: A Matter of Choices Made or Choices Given?” *Representation* 53(2): 153–66.

———. 2017b. “The Predictive Power of the Left-Right Self-Placement Scale for the Policy Positions of Voters and Parties.” *West European Politics* 40(2): 357–77.

Lupu, Noam, and Zach Warner. 2017. “Mass–Elite Congruence and Representation in Argentina.” In *Malaise in Representation in Latin American Countries*, eds. Alfredo Joignant, Mauricio Morales, and Claudio Fuentes. New York: Palgrave Macmillan US, 281–302. http://link.springer.com/10.1057/978-1-137-59955-1\_12 (February 14, 2020).

Milburn, Michael A. 1987. “Ideological Self-Schemata and Schematically Induced Attitude Consistency.” *Journal of Experimental Social Psychology* 23(5): 383–98.

Miller, Warren E., and Donald E. Stokes. 1963. “Constituency Influence in Congress.” *American Political Science Review* 57(01): 45–56.

Radelet, Steven, Michael Clemens, and Rikhil Bhavnani. 2005. “Aid and Growth.” *Finance and Development* 42(3): 1–8.

Romeijn, Jeroen. 2018. “Do Political Parties Listen to the(Ir) Public? Public Opinion–Party Linkage on Specific Policy Issues.” *Party Politics*: 135406881878734.

Rosset, Jan, and Christian Stecker. 2019. “How Well Are Citizens Represented by Their Governments? Issue Congruence and Inequality in Europe.” *European Political Science Review* 11(02): 145–60.

Ruedin, Didier. 2012. “Individual Representation: A Different Approach to Politial Representation.” *Representation* 48(1): 115–29.

Schakel, Wouter, and Armen Hakhverdian. 2018a. “Ideological Congruence and Socio-Economic Inequality.” *European Political Science Review* 10(3): 441–65.

———. 2018b. “Ideological Congruence and Socio-Economic Inequality.” *European Political Science Review* 10(3): 441–65.

Stimson, James A., Michael B. Mackuen, and Robert S. Erikson. 1995. “Dynamic Representation.” *American Political Science Review* 89(3): 543–65.

Sturgis, Patrick, and Patten Smith. 2010. “Fictitious Issues Revisited: Political Interest, Knowledge and the Generation of Nonattitudes.” *Political Studies* 58(1): 66–84.

Thomassen, Jacques. 1994. “Empirical Research into Political Representation: Failing Democracy or Failing Models?” In *Elections at Home and Abroad: Essays in Honor of Warren Miller*, eds. M. Kent Jennings and Thomas E. Mann. Ann Arbor: The University of Michigan Press, 237–65.

———. 2012. “The Blind Corner of Political Representation.” *Representation* 48(1): 13–27.

Todosijević, Bojan. 2004. “The Hungarian Voter: Left–Right Dimension as a Clue to Policy Preferences.” *International Political Science Review* 25(4): 411–33.

Wagner, Markus, and Outi Ruusuvirta. 2011. “Matching Voters to Parties: Voting Advice Applications and Models of Party Choice.” *Acta Politica* 47(4): 400–422.

Walczak, Agnieszka, and Wouter van der Brug. 2013. “Representation in the European Parliament: Factors Affecting the Attitude Congruence of Voters and Candidates in the EP Elections.” *European Union Politics* 14(1): 3–22.

Walgrave, Stefaan, and Jonas Lefevere. 2013. “Ideology, Salience, and Complexity: Determinants of Policy Issue Incongruence between Voters and Parties.” *Journal of Elections, Public Opinion and Parties* 23(4): 456–83.

Walgrave, Stefaan, and Christophe Lesschaeve. 2017. “The Matthew Effect in Electoral Campaigns: Increasing Policy Congruence Inequality during the Campaign.” *Electoral Studies* 50: 1–14.

———. 2018. “Policy and Ideology Volatility During the Campaign.” In *Mind the Gap. Political Participation and Representation in Belgium*, ed. Kris Deschouwer. Colchester: ECPR Press.

Werner, Annika. 2020. “Representation in Western Europe: Connecting Party-Voter Congruence and Party Goals.” *The British Journal of Politics and International Relations* 22(1): 122–42.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge [England] ; New York, NY, USA: Cambridge University Press.

Table 1: Analyses of OCV

|  |  |  |
| --- | --- | --- |
|   | Model 1 | Model 2 |
|   | B | S.E. | Sig. | B | S.E. | Sig. |
| Number of policy statements (cube root) | -7.92 | 0.02 | \*\*\* | -7.91 | 0.02 | \*\*\* |
| Topic diversity (cube root) | -4.69 | 0.02 | \*\*\* | -4.70 | 0.02 | \*\*\* |
| Personal salience | 0.47 | 0.02 | \*\*\* | 0.47 | 0.02 | \*\*\* |
| Political interest | -0.16 | 0.06 | \*\* | -0.61 | 0.07 | \*\*\* |
| Left-right statements | -0.46 | 0.01 | \*\*\* | -0.79 | 0.03 | \*\*\* |
| Policy domain difference | -0.28 | 0.09 | \*\* | -0.28 | 0.01 | \*\*\* |
| Political interest\*Left-right statements |  |  |   | 0.05 | 0.01 | \*\*\* |
| Lower education (ref. cat.) |  |  |   |  |  |  |
| Middle education | -0.19 | 0.42 |   | -0.21 | 0.42 |  |
| Higher education | -0.68 | 0.43 |   | -0.69 | 0.43 |  |
| Gender | -0.26 | 0.27 |   | -0.26 | 0.27 |  |
| Age | -0.01 | 0.01 |   | -0.01 | 0.01 |  |
| Income | -0.01 | 0.06 |   | -0.01 | 0.06 |  |
| Constant | 64.59 | 0.86 | \*\*\* | 67.31 | 0.89 | \*\*\* |
| Part controls | Yes | Yes |
| n (total/voters) | 9,681,030/844 | 9,681,030/844 |
| ∆ AIC | -2,422,587 | -2,422,683 |
| Note: \* = p ≤ 0.05; \*\* = p ≤ 0.01; \*\*\* = p ≤ 0.001 |

Table 2: Analyses of the correctness of voters’ party preferences

|  |  |  |
| --- | --- | --- |
|   | Model 1 | Model 2 |
|   | B | S.E. | Sig. | B | S.E. | Sig. |
| Number of policy statements (cube root) | 5.53 | 0.15 | \*\*\* | 5.60 | 0.22 | \*\*\* |
| Topic diversity (cube root) | 3.87 | 0.07 | \*\*\* | 2.84 | 0.14 | \*\*\* |
| Personal salience | -0.32 | 0.09 | \*\*\* | -0.11 | 0.20 |  |
| Political interest | 0.27 | 0.18 |  | 0.34 | 0.18 |  |
| Left-right statements | 1.39 | 0.10 | \*\*\* | 1.52 | 0.10 | \*\*\* |
| Policy domain difference | -9.34 | 0.71 | \*\*\* | -8.94 | 0.72 | \*\*\* |
| Lower education (ref. cat.) |  |  |  |  |  |  |
| Middle education | 1.25 | 1.41 |  | -3.02 | 1.51 | \* |
| Higher education | 2.92 | 1.43 | \* | -0.76 | 1.54 |  |
| Gender | 0.92 | 0.87 |  | 0.83 | 0.88 |  |
| Age | -0.02 | 0.03 |  | -0.02 | 0.03 |  |
| Income | 0.39 | 0.19 | \* | 0.42 | 0.19 | \* |
| Number of policy statements (cube root)\*Lower education (ref. cat.) |  |  |  |  |  |  |
| Number of policy statements (cube root)\*Middle education |  |  |  | 0.67 | 0.21 | \*\* |
| Number of policy statements (cube root)\*Higher education |  |  |  | -0.30 | 0.21 |  |
| Topic diversity (cube root)\*Lower education (ref. cat.) |  |  |  |  |  |  |
| Topic diversity (cube root)\*Middle education |  |  |  | 0.85 | 0.17 | \*\*\* |
| Topic diversity (cube root)\*Higher education |  |  |  | 1.19 | 0.17 | \*\*\* |
| Personal salience\*Lower education (ref. cat.) |  |  |  |  |  |  |
| Personal salience\*Middle education |  |  |  | -0.35 | 0.25 |  |
| Personal salience\*Higher education |  |  |  | -0.09 | 0.25 |  |
| Constant | 35.81 | 3.08 | \*\*\* | 37.54 | 3.11 | \*\*\* |
| n (total/voters) | 1,408,950/685  |  1,408,950/685  |
| ∆ AIC |  -234,514 |  -295,953 |
| Note: \* = p ≤ 0.05; \*\* = p ≤ 0.01; \*\*\* = p ≤ 0.001 |

Figure 1: Overview of the simulation process

List of *N* 134 statements

Step 1: draw sample of *n* statements from *N* and calculate congruence using Formula 1

Step 2: draw sample of *n* statements from *N-n* and calculate congruence using Formula 1

Step 3: calculate opinion congruence variability using Formula 3.

Step 4: Repeat step 1 to 3

Step 5: Repeat step 1 to 4 while increasing *n* to 66.

Figure 2: Marginal effects of the relation between the number of policy statements and OCV

 

Figure 3: Marginal effects of the relation between the statement sample size and OCV through segmented multilevel regression



Figure 4: Marginal effects of the relation between topic diversity and OCV

 

Figure 5: Density of the OCV distribution for the weighted and unweighted calculation.



Figure 6: Conditionality of the difference between high and politically interested voters on the number of left-right policy statements



Figure 7: Correct voting, education level, and topic diversity



1. We acknowledge that many studies examine opinion congruence on specific issues rather than congruence across a wide range of issues. While obviously more limited, the findings presented here should still be of interest, such as those regarding the number of policy statements. In addition, many scholarly endeavors remain concerned with a global assessments of the opinion congruence between voters and elites in their research questions (Romeijn 2018; Schakel and Hakhverdian 2018b; Walgrave and Lesschaeve 2017; Werner 2020). It is primarily in the efforts to make those global measurements better that the contributions of this article should be viewed. [↑](#footnote-ref-1)
2. This definition thus excludes issues on which there is a broad social consensus. Including them would make the actual level opinion congruence by definition very high and render it an almost useless concept. [↑](#footnote-ref-2)
3. See Table A1 in the Appendix for a comparison between the sample and Flemish population. [↑](#footnote-ref-3)
4. While these efforts were largely successful, some statements inevitably touched upon more divisive issues than others. Therefore, we ran two robustness checks. In a first, we limited our analyses to congruence scores calculated on the basis of only those statements where at least two out of six parties had a different opinion than the other four. A second check limited the statements to those where the majority opinion among voters (in our sample) did not exceed 75%. Both analyses, reported in Tables A10 and A11 confirm the findings reported in the main analyses. [↑](#footnote-ref-4)
5. To test whether the survey length and the large number of statements lead to survey fatigue despite the two wave design, we repeated the analyses for only the policy statements asked in the first wave. The results, reported in Table A3 in the Appendix, are substantially identical. [↑](#footnote-ref-5)
6. We acknowledge that this is but one of several conceptualizations of voter-party opinion congruence (Golder and Stramski 2010; Ruedin 2012), though it is the one most frequently encountered in the literature. Extending the analyses to other conceptualizations is, however, beyond the scope of this study. [↑](#footnote-ref-6)
7. All codings were done by the authors. A second coding indicated this coding to be reliable (Kippendorf’s Alpha of 0.83). [↑](#footnote-ref-7)
8. The maximum possible value is the size of the statement sample times two. [↑](#footnote-ref-8)
9. To assess the robustness of the findings presented in Table 1, we conducted three additional analyses. First, we tested whether the results hold when OCV is aggregated across the entire sample of voters. The results, reported in the Appendix in Table A7, confirm the findings of Table 1. In addition, we tested whether similar results are found when we limit the analyses to statements belonging to one ideological dimension (economic or cultural). This proved to be the case, as is reported in Table A8 in the Appendix. Third, we tested and confirmed in Table A9 in the Appendix that the results hold the analyses are run for each of the six political parties separately. [↑](#footnote-ref-9)
10. These analyses are performed on only the opinion congruence estimates of the first statement sample (step 1 in Figure 1). However, the results are similar if the estimates of the second sample are used. [↑](#footnote-ref-10)