European Journal of Political Research **64**: 2146–2160, 2025
doi: 10.1111/1475-6765.70022

Establishing the construct and predictive validity of brief measures of affective polarization

JAKOB KASPER,[1] GIJS SCHUMACHER[2] & BERT N. BAKKER[1]
[1]*Amsterdam School of Communication Research, University of Amsterdam, The Netherlands;* [2]*Department of Political Science, University of Amsterdam, The Netherlands*

Abstract. Measuring affective polarization, defined as the liking for one's political ingroup and the dislike for political outgroups, poses methodological challenges in multiparty systems: evaluations of seven, 13 or even more parties in a survey are costly, time-consuming and demanding. Some studies therefore use subsets of parties to create brief affective polarization measures. However, it is unclear how this affects the construct and predictive validity of these brief measures, potentially causing problematic inferences. Across 39 countries ($N = 66,880$), we demonstrate that brief measures that include ratings of only three to five parties can maintain acceptable validity, as illustrated by strong correlations with full measures and consistent associations with political correlates. The construct and predictive validity of brief measures are best when selecting a set of large, ideologically diverse parties. We provide specific recommendations for the effective measurement of affective polarization in different multiparty systems.

Keywords: affective polarization; brief measure; construct validity; predictive validity

Measuring affective polarization in multiparty systems

The extent and consequences of affective polarization capture both public (e.g., Applebaum, 2018; Edsall, 2023) and academic interest (e.g., Bakker & Lelkes, 2024; Hobolt et al., 2023; Iyengar et al., 2019). Affective polarization is the degree to which citizens are sympathetic to their political ingroup and antagonistic to political outgroups (Wagner, 2021). A common way to measure affective polarization is by asking participants how cold or warm they feel towards each political party in their country (Druckman & Levendusky, 2019; Lelkes, 2016). This is a straightforward and quick task in a two-party system such as the United States. But what if a party system includes seven, 13 or even 19 parties as is the case in some multiparty systems? Is it necessary to collect ratings for all parties? Are brief measures covering a subset of parties sufficient, and if so, which parties should be selected? So far, there is no established method to solve this problem – a gap this paper seeks to fill.

Having to evaluate feelings towards a large set of political parties in a survey can be problematic for various reasons. First, it may lead to survey fatigue (Galesic & Bosnjak, 2009): Participants may drop out of surveys (Edwards et al., 2004) or lose attention, which compromises the reliability of measures (Burisch, 1984). This is particularly concerning for hard-to-reach populations such as political elites, journalists or adolescents. Second, it is costly to include many items to measure a single concept, and omnibus surveys, such as national election studies, measure many different concepts, limiting the number of items one can include for each one. Third, in a rapidly evolving political landscape with numerous parties, citizens may have difficulties keeping up with new developments and forming clear opinions on all parties (Somin, 2018). Therefore, reducing the

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

number of parties assessed in surveys would reduce the burden on participants while freeing up space to measure additional concepts.

Scholars have already recognized this problem. For instance, Gidron et al. (2022, p. 3) describe that, in a study in Israel (nine parties in parliament at the time of the study) they '*focus on a subset of five parties in order to reduce response fatigue'*. Specifically, they '*chose parties that represent the full breadth of the Israeli political spectrum.*'' Another example is the Comparative Study of Electoral Systems (CSES) – a collection of election studies and frequently used source to study affective polarization – that collects feelings towards a maximum of nine political parties based on vote shares in the most recent elections (for other recent examples, see Comellas & Torcal, 2023; Riera & Garmendia Madariaga, 2023; Reiljan et al., 2024). We refer to this restriction on a subset of political parties as 'brief affective polarization measures'.

Currently, there is no established, empirically grounded way of deriving such brief measures. This is problematic because the brief measures that are used may not be valid measures of affective polarization. Specifically, decreasing the number of items in a scale can affect the construct validity (i.e., the extent to which the measure reflects the construct one wishes to measure) and, relatedly, the predictive validity (i.e., the ability to explain variance in related variables) of the measurement (Bakker & Lelkes, 2018; McDermott, 2011). If the construct validity of measures of affective polarization is severely compromised by focusing on a subset of parties, this could lead to type M (magnitude) and type S (sign) errors when studying it as a source, consequence or correlate of other variables (Gelman & Carlin, 2014). A type M error would occur if associations between affective polarization and other measures differ substantially in magnitude, but not in direction when using a brief scale compared to the full scale. A type S error occurs when the sign of the association changes depending on the measure. Invalid measures of affective polarization could, therefore, severely affect the conclusions researchers draw about the causes and consequences of affective polarization. Note that our research does not focus on the reliability of the full affective polarization measure – often evaluated through internal consistency or temporal stability, as discussed, for example, by Carlin and Love (2022) – but rather on the validity of brief measures (Bakker & Lelkes, 2018; McDermott, 2011). It is beyond the scope of this research note to further assess the reliability of full or brief measures.

In this paper, we provide evidence for the construct and predictive validity of brief affective polarization measures across 39 multiparty systems. We illustrate that one can derive valid measures of affective polarization by covering as little as three to five political parties in national surveys (depending on the size of the electoral system), as long as the selected parties vary to some extent in ideology and cover a substantial share of the electorate.

Data sources: Dutch case study and replication across 39 countries

We draw on two different data sources: (A) the Dutch Longitudinal Internet Studies for the Social Sciences (LISS; Scherpenzeel & Das, 2010) and (B) the CSES (CSES module 5, 2023).

The LISS is a panel study with a sample representative of the Dutch adult population. We use the data on politics and values (December 2020–March 2021, Wave 13), including responses from $N = 5613$ participants. At the time of data collection, 13 parties were represented in the Dutch parliament, all of which were covered in the survey.

The CSES is a global collection of election studies frequently used in research on affective polarization (e.g., see Harteveld, Mendoza, et al., 2022; Reiljan et al., 2024). We use data from

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

Module 5 (2016–2021), selecting only countries where feelings towards at least five political parties were assessed, as there is less reason to further shorten measures of affective polarization based on evaluations of four or fewer parties. This selection process resulted in a final sample of $N_{ovr} = 61,267$ participants across 39 countries on six different continents (Online Appendix A5 describes sample sizes and data collection periods for each country[1]). The covered countries vary substantially in political histories, systems, population size and economic status.

We first focus on the Dutch case and then extend findings to a broader international context. The Netherlands' highly fragmented multi-party system offers an ideal case to test the validity of brief measures. Further, the LISS is more comprehensive in terms of the number of parties covered in the survey, since the CSES assesses feelings towards a maximum of nine parties not always including all parties represented in parliament. Extending our analysis to the CSES allows us to test the generalizability of our findings and provide country-level recommendations for diverse political contexts.

Measures and operationalization of construct and predictive validity

Affective polarization and construct validity

In the LISS panel, feelings towards the 13 political parties represented in the Dutch parliament were assessed with the question: '*How sympathetic do you find the political parties? You can assign each party a score between 0 and 10. 0 means that you find the party very unsympathetic, and 10 means that you find the party very sympathetic'*. In the CSES, the wording was '*After I read the name of a political party, please rate it on a scale from 0 to 10, where 0 means you strongly dislike that party and 10 means that you strongly like that party'*. As recommended by Wagner (2021) when introducing the applied measures of affective polarization, we categorized 'Don't know' responses as missing values. We operationalize affective polarization using the spread of like/dislike across the political spectrum, weighted by the relative vote share of each party, using the following equations (Wagner, 2021):

$$\overline{like}_i = \sum_{p=1}^{P} (v_p \cdot like_{ip}), \tag{1}$$

$$WSpread_i = \sqrt{\sum_{p=1}^{P} v_p \cdot (like_{ip} - \overline{like}_i)^2}. \tag{2}$$

Here, $\overline{like}_i$ represents the weighted mean like/dislike score for individual i, calculated by multiplying the like/dislike score for each party ($like_{ip}$) by its relative vote share (v_p). $WSpread_i$ is the weighted spread of like–dislike ratings for individual i across parties p, computed as the root mean square deviation of the like/dislike ratings for each party from an individual's weighted mean like/dislike rating, again considering the relative vote share of each party. Affective polarization is low when individuals exhibit uniform feelings towards all parties – whether positive, negative or neutral – and high when they display strong positive feelings towards specific parties, coupled with negative feelings towards others. By including vote shares, this formula prioritizes the influence of larger parties on an individual's affective polarization score, as it is arguably more consequential if the liked/disliked parties are larger competitors rather than smaller fringe parties (Wagner, 2021).

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

To operationalize construct validity, we compare affective polarization measures based on all political parties within each country (N_p) to affective polarization measures based on all possible combinations of 2 to $N_p - 1$ parties. For example, in the Netherlands, we calculated one full affective polarization score with all 13 parties and brief scores with all combinations of 2–12 parties (8178 scores in total per participant). We excluded participants only when the calculation of the score was infeasible (i.e., number of evaluated parties < 2), to avoid removing participants who did not rate lesser known parties. We then operationalize construct validity (i.e., the extent to which the full measures – as the current standard – and the brief measures capture the same construct) by correlating the brief affective polarization measures with the full measures derived from evaluations of all parties. We follow the recommendations of Nunnally and Bernstein (1994) and Ellis (2013) and assume that a construct validity of $r = 0.80$ or higher is acceptable. Adding more items beyond this point typically yields diminishing returns in terms of increasing construct validity and precision of associations with other variables.

Note that we replicate our results with three alternative affective polarization measures, more specifically, the unweighted spread of like/dislike and the weighted and unweighted mean distance of liking from the favorite party – the latter being more consistent with conceptualizations of affective polarization as revolving mainly around a single political ingroup (Wagner, 2021). The results align well. We provide detailed findings in the online Appendices (see Sections 1 and 6 for unweighted spread, Sections 2 and 8 for unweighted mean distance, and Sections 4 and 11 for weighted mean distance).

Political correlates and predictive validity

To evaluate the predictive validity of the brief measures, we calculated bivariate correlations with often studied correlates of affective polarizations (Torcal & Harteveld, 2025; Wagner, 2024). Specifically, we assess the relationship between affective polarization and satisfaction with democracy, confidence in democracy, party identification, partisan identity strength, political interest and turnout intentions in the LISS and associations with satisfaction with democracy, previous turnout, perceived efficacy of voting, political interest, partisan identity strength and political identification in the CSES. Item formulations can be found in Online Appendices 1.3.1 (LISS) and 6.2.1 (CSES). While the selected political correlates are not an exhaustive representation of all constructs that have been associated with affective polarization, nor are they exclusively related to affective polarization, we focus on them due to their frequent use in recent studies on affective polarization (e.g., see Harteveld & Wagner, 2023; Janssen & Turkenburg, 2024; Wagner, 2021).

When evaluating the predictive validity of brief measures, we compare the associations between brief measures and political correlates with those observed with the full measures. Thereby, we examine whether brief measures yield results comparable to the current standard measure in terms of relationships with theoretically relevant constructs.

Results

The construct validity of brief measures

The upper panel of Figure 1 shows distributions, means and 95% confidence intervals (CIs) of the correlations between brief and full affective polarization measures (LISS, the Netherlands)

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

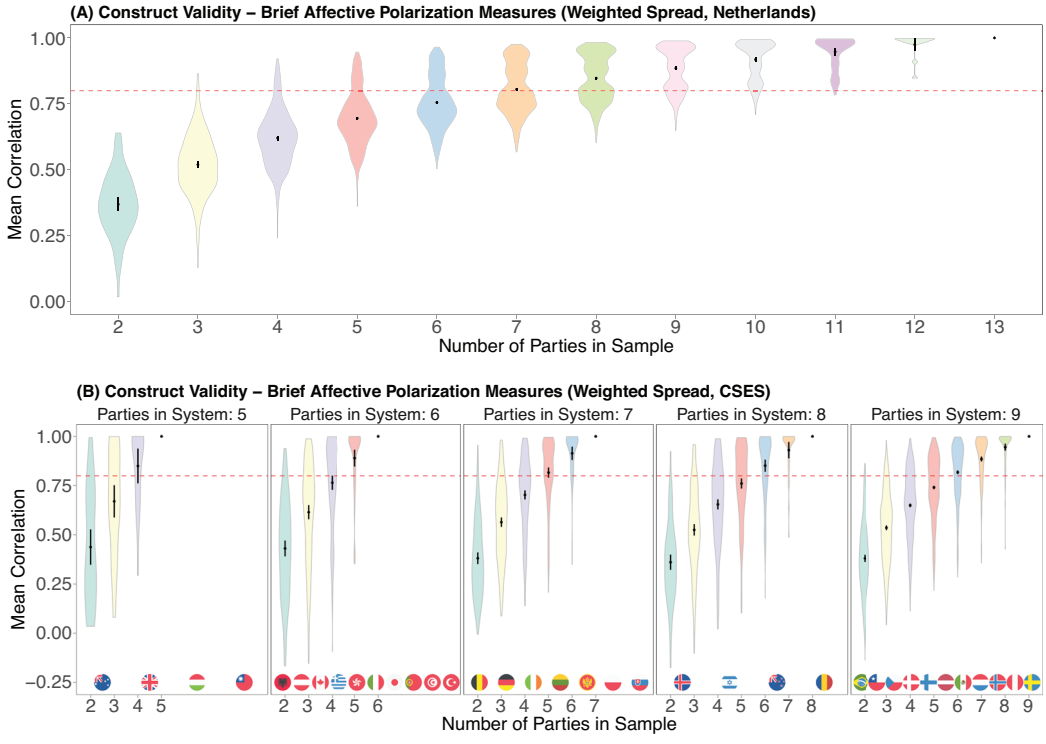


Figure 1. Construct validity – brief affective polarization measures (weighted spread).
Note: Distributions and means of correlations between the full and brief affective polarization scores (weighted spread of like/dislike) with 95% CIs in the Netherlands (LISS, panel A), and across 39 countries (CSES, panel B). The red-dashed line in both panels represents a correlation of $r = 0.80$, which was defined as the threshold for acceptable construct validity. Panel B is divided based on the highest number of parties included in the full affective polarization measure as per availability in the CSES. Flags indicate which countries are included in which sub-panel – they are equally distributed in the sub-panels and do not correspond to specific data points.

across different numbers of parties included in the sample for the calculation of the brief measures (see Online Appendix 3.2 for more details). While measures that include only three parties almost always fail to reach the aspired construct validity of $r = 0.80$ ($M_r = 0.52$, $SD_r = 0.11$, $range_r[0.13; 0.87]$), brief measures with four parties already include a larger set of correlations above the specified threshold ($M_r = 0.62$, $SD_r = 0.11$, $range_r[0.24; 0.92]$). For brief measures that include five parties, the mean correlation with the full measure is $M = 0.70$ ($SD_r = 0.10$, $range_r[0.36; 0.95]$), with many correlations surpassing $r = 0.80$. Including additional parties raises the construct validity even more consistently over the specified threshold.

The bottom panel of Figure 1 summarizes the construct validity of the brief affective polarization measures for the countries included in the CSES. Party system size always refers to the highest number of parties available in the dataset for each country.[2] While measures that include only three parties seem to produce an extensive set of combinations that reach the aspired construct validity for party systems with five parties ($M_r = 0.67$, $SD_r = 0.26$, $range_r[0.08; 1.00]$) or six parties ($M_r = 0.61$, $SD_r = 0.26$, $range_r[-0.15; 0.99]$), combinations with four parties more consistently provide valid measures for systems with seven parties ($M_r = 0.70$, $SD_r = 0.19$,

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

$range_r$[0.14; 0.99]), eight parties ($M_r = 0.65$, $SD_r = 0.21$, $range_r$[0.02; 0.99]) or nine parties ($M_r = 0.65$, $SD_r = 0.17$, $range_r$[0.11; 0.99]). Including more parties in the measures within each system led to an even more consistent crossing of the threshold of $r = 0.80$.

In both the LISS and the CSES, we found combinations with a specific number of parties that consistently achieve an acceptable construct validity (e.g., when reducing the number of parties by just one or two). However, briefer versions of the scales (e.g., four/five parties in the LISS and three/four in the CSES), despite performing worse on average, still have a notable proportion of combinations that exceed the specified validity threshold. On the flip side, while combinations of larger numbers of parties are more consistent in achieving sufficient construct validity, they are not always a safeguard against measures with low validity, as evidenced by the wide range of correlations within the same scale length. We conclude that while *mean* correlations between brief affective polarization measures may be lower or higher than the threshold of $r = 0.80$, *individual* brief measures of the *same length* with *specific* party combinations can vary from being nearly identical to the full measure ($r > 0.90$) to exhibiting very low associations. This raises the question: Is it possible to predict which party combinations are likely to maintain an acceptable validity and if so, how can we identify the most effective combinations of parties to maximize construct validity while minimizing scale length?

To answer this question, we examine the composition of combinations of parties that can be selected to likely attain acceptable construct validity focusing on party size and ideological diversity. Those are the two aspects that have been considered by other researchers who have relied on subsets of parties to measure affective polarization (e.g., Comellas & Torcal, 2023; Gidron et al., 2019). However, again, the approach has never been formally evaluated. To operationalize ideological diversity, we use the latest version of the Chapel Hill Expert Survey (CHESS; Jolly et al., 2022) from 2019 for the LISS, and comparable expert assessments available in the CSES. Both provide estimates of parties' ideological positions on a scale from 0 (*left*) to 10 (*right*)[3]. We calculate mean ideological diversity as the mean pairwise difference between the ideological positions of parties within each combination of parties. Additionally, we operationalize the mean party size for each combination of parties as the average vote share of all parties within a given combination, based on the most recent lower house elections.

To examine the relationship between construct validity and both characteristics in the Netherlands (LISS), we employed a linear regression with the correlation between the brief and full measures of affective polarization as the dependent variable, and mean ideological diversity and mean party size within the combination of parties (both z-standardized, i.e., $M = 0$, $SD = 1$) and their interaction as predictors. We control for the number of parties included in the measure. We find a statistically significant positive association between mean ideological diversity and the correlation between the full and brief affective polarization measures ($\beta = 0.024$, 95% CI [0.023,0.025], $p < 0.001$). At the mean level of party size, combinations covering a broader ideological spectrum achieve a higher construct validity (i.e., a 2*SD* increase in ideological diversity would amount to an increase in the Pearson correlation between the full and brief measure by about 0.05). Additionally, we find a statistically significant positive association with mean party size ($\beta = 0.067$, 95% CI [0.066,0.068], $p < 0.001$), indicating that at the mean level of ideological diversity, the party size within a combination is positively related to its construct validity (i.e., a 2*SD* increase would roughly correspond to a 0.14 increase in construct validity).

Further, a statistically significant interaction between mean ideological diversity and mean party size, albeit with a very small effect size, suggests that the positive associations of ideological

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

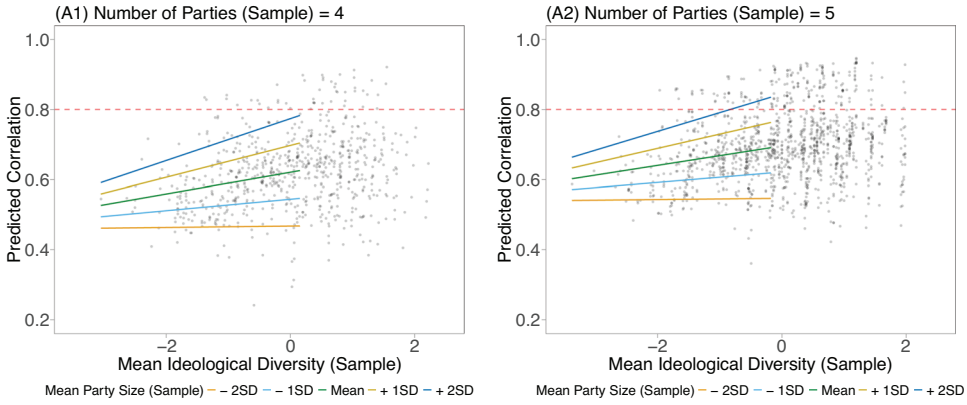


Figure 2. Construct validity by ideological diversity and party size (weighted spread, NL).
Note: Panels A1 and A2 plot the predicted correlation between the full and brief affective polarization scores as a function of the mean ideological diversity and mean party size (+2 SD, +1 SD, mean, −1 SD, −2 SD) within samples of four and five parties (both z-standardized), respectively, used to calculate the brief scores of affective polarization. Both predictors were z-standardized. The red-dashed line represents a correlation of $r = 0.80$.

diversity and party size with construct validity are enhanced in combinations with larger average party size and more ideological diversity, respectively ($\beta = 0.006$, 95% CI [0.005,0.007], $p < 0.001$). Figure 2 plots the predicted correlation between full and brief affective polarization measures as a function of mean ideological diversity and mean party size for combinations of 4 (Panel A1) and 5 (Panel A2) parties. To illustrate the metric, consider the following examples of four-party combinations: While the party combination with the highest construct validity (VVD, PVV, GL, SP; $r_{construct} = 0.92$) has a considerable ideological diversity (*Standardized mean pairwise differences* $= 1.55$) and average party size (*Mean vote share* $= 0.13$), the party combination with the lowest construct validity (PVV, FvD, SGP, DENK; $r_{construct} = 0.24$) contains small (*Mean vote share* $= 0.05$), ideologically homogeneous (*Standardized mean pairwise differences* $= -0.59$) parties. Online Appendix A2 lists additional party combinations with high/low construct validity.

To assess the relationship between construct validity and mean ideological diversity and party size in the CSES, we employed a linear mixed model with the correlation between the full and brief affective polarization measures as the criterion, and the z-standardized mean party size (i.e., vote share) and mean ideological diversity (i.e., pairwise ideological differences) of parties within a combination as well as their interaction as predictors. We also controlled for the number of parties included in the brief measures and, as the estimated variance for an additional random effect of the party system was close to zero, we used a simplified model that includes only a random intercept for countries. We find statistically significant positive associations between construct validity and mean ideological diversity ($\beta = 0.046$, 95% CI [0.043,0.049], $p < 0.001$) and mean party size ($\beta = 0.152$, 95% CI [0.149,0.154], $p < 0.001$). Given the multiplicative structure of our model, this indicates that at the mean level of party size and ideological diversity, respectively, both variables are positively related to construct validity. Further, the interaction effect suggests that those relationships are not conditional on the other variable, as the interaction term is very close to zero and not statistically significant ($\beta = 0.001$, 95% CI [−0.002,0.003], $p = 0.560$).

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

The positive association between construct validity and both ideological diversity and party size is also reflected in the *'optimal'* party combinations for each country that mostly encompass large parties that cover the whole ideological spectrum. For instance, for Germany the ideal combination contains ratings for the Social Democratic Party (SPD), the Christian Democratic Union (CDU), the Green Party (Gruene), and the Alternative for Germany (AFD) ($r = 0.93$), while for Denmark it encompassed the Social Democrats (SD), Venstre (Liberal Party, V), the Danish People's Party (DF) and the Red-Green Alliance (EL) ($r = 0.92$). It is noteworthy that, despite the use of different data sources, the '*optimal*' four-party combination determined with the LISS data in the Netherlands (VVD, PVV, SP, GL) very closely matches the combination obtained in the CSES (VVD, PVV, SP, D66, $r = 0.93$). Three out of four parties are identical and both combinations contain another progressive party as the fourth party. Further attesting to the robustness of our results, the combination identified with the LISS also exhibited high construct validity when evaluated with CSES data ($r = 0.86$). In Online Appendix A6, we provide the optimal combinations for each of the studied countries.

Across 39 countries, we demonstrate that affective polarization measures derived from feelings towards three to five parties can have acceptable construct validity when they consist of large parties with sufficient ideological diversity. Our main findings and even specific party recommendations remain robust for alternative conceptualizations of affective polarization (i.e., unweighted spread, unweighted/weighted mean distance; Wagner, 2021). Detailed results are described in Online Appendices 1, 2, 4, 6, 8 and 11. We conclude that even with a substantial reduction in the number of parties, brief measures still capture what we intend to capture with the full affective polarization measures. In the next section, we will examine how shortening scales influences associations with frequently studied political correlates.

The predictive validity of brief measures

We assess predictive validity by examining bivariate correlations between the full/brief affective polarization measures and a set of often studied correlates. The correlates are scored such that positive correlations indicate that an increase in affective polarization is associated with an increase in the target variables. While we examine frequently studied *outcomes* of affective polarization, our design – like many other studies in this literature (e.g., Berntzen et al., 2024; Gidron et al., 2022; Wagner, 2021) – does not allow for direct causal inferences.

We plot the mean correlations between the political correlates and full/brief affective polarization measures in the Netherlands (LISS) in panel A1 of Figure 3 (see Online Appendix 3.3 for more details). Consistent with existing literature (Harteveld, Berntzen, et al., 2022; Harteveld & Wagner, 2023; Wagner, 2021), all correlations with affective polarization were positive.[4] Similar to the results for construct validity, the mean correlations remain relatively stable when different numbers of parties are included in the brief affective polarization measures. Notably, the truncation of the scores did not result in Type S errors when examining mean correlations as the signs of the correlations are the same when using brief or full measures. However, there are indications of Type M errors. For some constructs (e.g., turnout, identity strength), we see considerable declines in the correlations when comparing the full measure of affective polarization to brief measures based on four or five parties. For turnout, for instance, the full affective polarization score correlates at $r = 0.28$, while mean correlations with measures derived from four or five parties are $M_r = 0.19$ and $M_r = 0.21$, respectively.

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

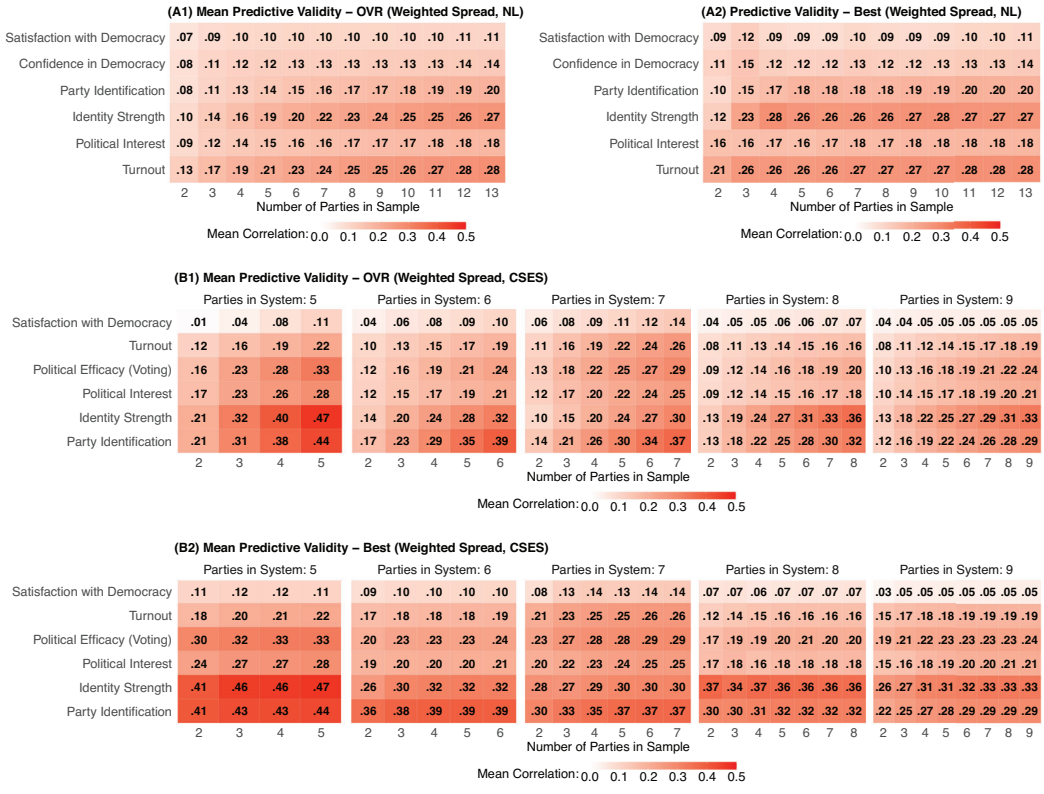


Figure 3. Predictive validity – brief affective polarization measures (weighted spread).
Note: Bivariate correlations between the full and brief affective polarization measures (weighted spread of like/dislike) and frequently studied political correlates. Panels A1 and B1 show the mean correlations for all possible combinations of a given number of parties in the Netherlands (LISS, A1) and across 39 countries (CSES, B1). Panel A2 shows the correlations for the optimal party combinations for a given scale length in the Netherlands (LISS), selected based on construct validity. Panel B2 shows the mean correlations for the optimal party combinations for a given scale length (based on construct validity) across the 39 countries in the CSES dataset. Panels B1 and B2 are divided based on the highest number of parties included in the full affective polarization measures as per the availability for each country in the CSES data. Further details (e.g. *SE*, 95% CIs) can be found in Online Appendices 3.3. and 9.2.

In Figure 1, we demonstrated substantial variability in the construct validity of brief affective polarization measures of the same length. This implies that associations with other variables may also not be uniform across brief measures of the same length that are based on different party combinations. Looking at four-party combinations, we observe a wide range of correlations with satisfaction with democracy ($range_r[-0.23; 0.27]$), confidence in democracy ($range_r[-0.22; 0.28]$), party identification ($range_r[0.01; 0.20]$), identity strength ($range_r[0.02; 0.29]$), political interest ($range_r[0.01, 0.21]$) and turnout ($range_r[0.04; 0.28]$). It is apparent that, despite stable *mean* correlations across the brief affective polarization measures, the application of *individual* brief measures with *specific* party combinations can lead to notable variations in magnitude (Type M error) and even sign (Type S error) of associations. This is unsurprising given that brief measures calculated from representative, ideologically diverse sets

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

of parties seem to mirror full measures in capturing affective polarization (i.e., high construct validity), while measures focusing on small, uniform parties may reflect different constructs (i.e., low construct validity) and, hence, associations.

To ascertain whether party combinations with high construct validity also yield an enhanced predictive validity, we examine the correlations between the '*optimal*' party combinations of a given length, as determined by their construct validity, and the political correlates in panel A2 of Figure 3. The results demonstrate remarkable consistency between the full and '*optimal*' brief measures in their associations with political correlates across the entire length of the scale, with only minimal variation. This emphasizes that by carefully selecting party combinations for construct validity, we can also achieve a good predictive validity with brief affective polarization measures.

We repeated this approach across the 39 countries included in the CSES data. Mean correlations across different party systems and numbers of parties included in the brief affective polarization measures are plotted in panel B1 of Figure 3. Again, all correlations with affective polarization were positive. In each party system, we find considerable stability of the mean correlations across the different numbers of parties used to calculate affective polarization, up until a certain number of included parties that aligns well with the previously identified points at which the mean construct validity begins to decline (i.e., three parties in five- or six-party systems, and four parties in seven-, eight- or nine-party systems). However, again, we find substantial variations in the correlations between brief affective polarization measures and target variables across different combinations of the same number of parties within a country (see Online Appendix 9.2). Thus, we again explored if selecting combinations based on high construct validity also improves the predictive validity. To assess this, we selected the party combinations with the highest construct validity for each length of the brief affective polarization measures within each country. Then, we calculated the mean correlations of these optimal combinations with the target variables for each length of the affective polarization measure within each party system. The results are displayed in panel B2 of Figure 3. Again, the correlations are strikingly consistent across the full measures and *'optimal'* brief affective polarization measures and there is only minimal variation across the length of the brief measures.

In line with our findings on construct validity, we observe that brief affective polarization measures, based on feelings towards only three to five political parties, can exhibit acceptable predictive validity when the parties are selected based on high construct validity. This alignment between findings for both metrics also suggests that there is no trade-off between achieving higher construct validity and maintaining strong predictive validity. We find this consistently across both the LISS and CSES datasets as well as across four different operationalizations of affective polarization (details on other measures are provided in Online Appendices 1.3, 2.3, 4.3, 6.2, 8.2 and 11.2).

Conclusion

Evidence from 39 multiparty systems ($N = 66,880$) demonstrates that scholars can rely on brief affective polarization measures without compromising the construct or predictive validity of their measurement. Depending on the number of parties in the political system, assessing participants' feelings towards three to five parties is sufficient. Yet, maintaining high validity levels when using brief measures requires selecting an ideologically diverse set of large parties

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

(e.g., including major parties from different ideological camps). Our extensive cross-national replication demonstrates that our findings are robust across a wide range of political contexts. Moreover, the results were consistent across all four metrics of affective polarization examined. For each country, we provide specific recommendations for the optimal combination of parties to include in brief measures of affective polarization in the Online Appendix (Sections 3.2.4 and 9.1.3). To facilitate the application of our findings, we developed an interactive Shiny App (https://jakob-kasper.shinyapps.io/Brief_AP_Measures_Kasper_et_al_2025/). This tool allows users to explore the construct validity of each affective polarization measure for different combinations and numbers of parties across the included countries, based on Wave 5 of the CSES data. The app will be updated to incorporate Wave 6 data upon its full release. We want to emphasize that the current recommended sets of parties are based on the data available in each country at the time of this study. Going forward, scholars who wish to use brief affective polarization measures should always carefully validate their brief measures using the most recent data. To do so, we encourage researchers to complement the outlined heuristic and recommendations with a data-driven approach. Our replication package contains code to evaluate all potential party combinations in any political context (see https://osf.io/rc9hg/?view_only=246d4330bbc147b5ac347cfb1f7b8f79).

Our results are in line with recent approaches of conceptualizing affective polarization through political camps or blocs, showing that complete party representation is not necessary for valid measures (Bantel, 2023; Kekkonen & Ylä-Anttila, 2021). Our approach allows researchers to shorten surveys – thereby reducing respondent fatigue, making room for additional constructs and focusing on known parties that are most relevant to citizens' level of polarization (and excluding parties that may be unknown to a majority of citizens) – while maintaining a valid measure of affective polarization *and* assessing potentially relevant party-level preferences. Thus, brief measures are not merely a compromise due to resource constraints but also offer distinct advantages.

Of course, our work leaves some open questions. First, are there other system-level differences that could influence the construct validity of brief measures of affective polarization? Our cross-national replication shows the robustness of our results across a wide range of political contexts, demonstrating that ideological diversity and party size are key factors in ensuring the validity of brief measures in the 39 analyzed countries. Nevertheless, other system-level differences (e.g., party structures, ideological polarization) might also affect the validity in specific countries. This underlines the importance of our recommendation to complement heuristics based on party size and ideological diversity with data-driven approaches, especially in countries not included in our analysis.

Second, can we compare absolute values of polarization between brief and full measures, or monitor absolute levels of affective polarization over time with brief measures? As explained above, we are interested in whether full and brief measures reflect the same construct and show consistent associations with other variables, rather than comparing absolute values (e.g., Boxell et al., 2024; Garzia et al., 2023), which may differ depending on scale length. If the goal is to observe longitudinal changes in *absolute levels* of affective polarization within a population, it is crucial to maintain consistent operationalizations in terms of scale length and the types of parties included in the measurement in order to make the values as comparable as possible. It should be noted, however, that this problem arises not only for brief affective polarization measures but also for full measures in longitudinal research – like the Dutch LISS panel we used in this study. For

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

instance, scales might change as the number of parties that are assessed in a survey increases or decreases, influential or newly formed parties are added, or parties that lose relevance are removed from the measurements in panel studies. We advise researchers aiming to make longitudinal comparisons to generally take such differences in the scales into account. The same applies to cross-country comparisons.

Third, do our results generalize to measures of horizontal polarization? In this article, we have focused on measures of affective polarization that assess participants' feelings towards political parties (i.e., *vertical* affective polarization) rather than their feelings towards voters of different political parties (i.e., *horizontal* affective polarization). We recognize the importance of extending the scope to horizontal measures, particularly when examining citizen dynamics. Although both measures correlate substantially (Druckman & Levendusky, 2019; Tichelbaecker et al., 2023), we encourage the integration of horizontal measures into national (panel) surveys such as those used here in order to further determine in future research whether and to what extent our conclusions can be generalized to measures of horizontal polarization.

To conclude, by systematically evaluating the quality of brief measures covering subsets of political parties, our research provides the much needed empirical backing for a practice that is necessary to efficiently operationalize affective polarization in multi-party systems. At the same time, our study cautions researchers to make informed choices about which parties to include in the applied measures, potentially utilizing publicly available datasets like the CSES or the LISS. In line with other studies (e.g., Bakker & Lelkes, 2018; Yeung & Quek, 2025), we highlight the necessity of validating key constructs in our research: a crucial step in generating research findings that can speak to complex political phenomena.

Acknowledgements

We would like to thank the members of the Hot Politics Lab, the IP-PAD doctoral network as well as Eelco Harteveld and Markus Wagner for their helpful comments on this work. Our work also benefited from the feedback received during presentations at Etmaal Communicatiewetenschap 2024 and the Hot Politics Lab meetings at the University of Amsterdam. We are also thankful to the reviewers at *EJPR* whose comments have significantly strengthened the manuscript. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement Nr. 101072992 (Interdisciplinary Perspectives on the Politics of Adolescence and Democracy). This publication is also part of the project 'Under pressure: How citizens respond to threats and adopt the attitudes and behaviours to counter them' [project number VI.Vidi.211.055 awarded to Bert N. Bakker] of the research programme NWO Talent Programme VIDI which is financed by the Dutch Research Council (NWO). Gijs Schumacher's research was supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme grant agreement No 759079 (POLEMIC).

Conflicts of interest statement

The authors declare no conflicts of interest in this research.

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

Ethics statement

The LISS panel, CSES and CHES data were not collected by the authors.

Data availability statement

Given the exploratory nature of our research, the studies were not pre-registered. Due to user agreements, the Dutch LISS panel, the CSES, and the CHES cannot be shared. On our OSF page, we provide a document 'Data Access' in which we explain how to access these datasets, the code to process the data and to reproduce all results, as well as the Online Appendix, see: https://osf.io/rc9hg/?view_only=246d4330bbc147b5ac347cfb1f7b8f79).

Online Appendix

Additional supporting information may be found in the Online Appendix section at the end of the article:

Data S1

Notes

1. The CSES data from the Netherlands consist of two components: One sample randomly drawn from population registers ($N = 1688$) and one sample drawn from the LISS ($N = 1797$). We remove participants sampled from the LISS to avoid analyzing the same data twice.
2. In total, there were 17 countries for which all parties represented in parliament at the time of data collection were addressed in the CSES (as was the case in the Dutch LISS panel), and 21 countries for which the CSES only assesses sentiments towards a subset of parties (see Online Appendix 5). Our conclusions are not affected by whether all or only a subset of parties were present in the CSES (see Online Appendix 10). For El Salvador and France, our analyses were limited to the unweighted spread of like/dislike and unweighted mean distance measure of affective polarization (see Online Appendices 6 and 8), as vote shares were unavailable. The results are substantively consistent with those presented here.
3. We use the left-right axis to operationalize ideological diversity, as it is most commonly used to distinguish the positions of different political parties in most of the countries studied here. In countries where other dimensions are more crucial or common (e.g., conservative vs. progressive), these alternative dimensions should be used to capture ideological diversity.
4. While previous studies investigating affective polarization and confidence in/satisfaction with democracy often report negative associations (Wagner, 2021), effects are usually small and inconsistent (Broockman et al., 2023). The reported differences may stem from our use of bivariate correlations and the conceptualization of affective polarization as spread of like/dislike rather than affective distance to a single ingroup.

References

Applebaum, A. (2018). *A warning from Europe: The worst is yet to come*. The Atlantic. https://www.theatlantic.com/magazine/archive/2018/10/poland-polarization/568324/?utm_source=copylink&utm_medium=social&utm_campaign=share

Bakker, B. N., & Lelkes, Y. (2018). Selling ourselves short? How abbreviated measures of personality change the way we think about personality and politics. *The Journal of Politics*, *80*(4), 1311–1325. https://doi.org/10.1086/698928

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

Bakker, B. N., & Lelkes, Y. (2024). Putting the affect into affective polarisation. *Cognition and Emotion*, *38*(4), 418–436. https://doi.org/10.1080/02699931.2024.2362366

Bantel, I. (2023). Camps, not just parties. The dynamic foundations of affective polarization in multi-party systems. *Electoral Studies*, *83*, 102614. https://doi.org/10.1016/j.electstud.2023.102614

Berntzen, L. E., Kelsall, H., & Harteveld, E. (2024). Consequences of affective polarization: Avoidance, intolerance and support for violence in the United Kingdom and Norway. *European Journal of Political Research*, *63*(4), 927–949. https://doi.org/10.1111/1475-6765.12623

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2024). Cross-country trends in affective polarization. *Review of Economics and Statistics*, *106*(2), 557–565. https://doi.org/10.1162/rest_a_01160

Broockman, D. E., Kalla, J. L., & Westwood, S. J. (2023). Does affective polarization undermine democratic norms or accountability? Maybe not. *American Journal of Political Science*, *67*(3), 808–828. https://doi.org/10.1111/ajps.12719

Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, *39*(3), 214. https://doi.org/10.1037/0003-066X.39.3.214

Carlin, R. E., & Love, G. J. (2022). Measuring affective polarization. In *Handbook of affective polarization* (pp. 79–102). https://files.osf.io/v1/resources/rh29g/providers/osfstorage/671eb6920dc87170666294aa?action=download&direct&version=4#page=79

Comellas, J. M., & Torcal, M. (2023). Ideological identity, issue-based ideology and bipolar affective polarization in multiparty systems: The cases of Argentina, Chile, Italy, Portugal and Spain. *Electoral Studies*, *83*, 102615. https://doi.org/10.1016/j.electstud.2023.102615

CSES Module 5. (2023). The Comparative Study of Electoral Systems. https://cses.org/data-download/cses-module-5-2016-2021/

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization? *Public Opinion Quarterly*, *83*(1), 114–122. https://doi.org/10.1093/poq/nfz0035

Edsall, T. B. (2023). Meet the people working on getting us to hate one another less. *The New York Times*. https://www.nytimes.com/2023/02/08/opinion/political-hatred-negative-partisanship.html

Edwards, P., Roberts, I., Sandercock, P., & Frost, C. (2004). Follow-up by mail in clinical trials: Does questionnaire length matter? *Controlled Clinical Trials*, *25*(1), 31–52. https://doi.org/10.1016/j.cct.2003.08.013

Ellis, J. L. (2013). A standard for test reliability in group research. *Behavior Research Methods*, *45*, 16–24. https://doi.org/10.3758/s13428-012-0223-z

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. https://doi.org/10.1093/poq/nfp031

Garzia, D., Ferreira da Silva, F., & Maye, S. (2023). Affective polarization in comparative and longitudinal perspective. *Public Opinion Quarterly*, *87*(1), 219–231. https://doi.org/10.1093/poq/nfad004

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Gidron, N., Adams, J., & Horne, W. (2019). Toward a comparative research agenda on affective polarization in mass publics. *APSA Comparative Politics Newsletter*, *29*, 30–36. https://ssrn.com/abstract=3391062

Gidron, N., Sheffer, L., & Mor, G. (2022). Validating the feeling thermometer as a measure of partisan affect in multi-party systems. *Electoral Studies*, *80*, 102542. https://doi.org/10.1016/j.electstud.2022.102542

Harteveld, E., Berntzen, L. E., Kokkonen, A., Kelsall, H., Linde, J., & Dahlberg, S. (2022). *The (alleged) consequences of affective polarization: Individual-level evidence & a survey experiment in 9 countries*. Preprint. https://osf.io/64uwd/

Harteveld, E., Mendoza, P., & Rooduijn, M. (2022). Affective polarization and the populist radical right: Creating the hating? *Government and Opposition*, *57*(4), 703–727. https://doi.org/10.1017/gov.2021.31

Harteveld, E., & Wagner, M. (2023). Does affective polarisation increase turnout? Evidence from Germany, the Netherlands and Spain. *West European Politics*, *46*(4), 732–759. https://doi.org/10.1080/01402382.2022.2087395

Hobolt, S., Lawall, K., & Tilley, J. (2023). The polarizing effect of partisan echo chambers. *American Political Science Review*. https://doi.org/10.1017/S0003055423001211

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.

Janssen, L., & Turkenburg, E. (2024). Breaking free from linear assumptions: Unravelling the relationship between affective polarization and democratic support. *European Journal of Political Research*. Advance online publication. https://doi.org/10.1111/1475-6765.12725

Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M., & Vachudova, M. A. (2022). Chapel Hill Expert Survey trend file, 1999-2019. *Electoral Studies*, *75*. https://doi.org/10.1016/j.electstud.2021.102420

Kekkonen, A., & Ylä-Anttila, T. (2021). Affective blocs: Understanding affective polarization in multiparty systems. *Electoral Studies*, *72*, 102367. https://doi.org/10.1016/j.electstud.2021.102367

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392–410. https://doi.org/10.1093/poq/nfw005

McDermott, R. (2011). Internal and external validity. In J. N. Druckman, D. P. Greene, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge handbook of experimental political science* (pp. 27–40). Cambridge University Press. https://www.cambridge.org/core/services/aop-cambridge-core/content/view/11F8092DCA6BECA1BBB922BFA08B4E23

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill.

Reiljan, A., Garzia, D., Ferreira da Silva, F., & Trechsel, A. H. (2024). Patterns of affective polarization toward parties and leaders across the democratic world. *American Political Science Review*, *118*(2), 654–670. https://doi.org/10.1017/S0003055423000485

Riera, P., & Garmendia Madariaga, A. (2023). Overlapping polarization: On the contextual determinants of the interplay between ideological and affective polarization. *Electoral Studies*, *84*, 102628. https://doi.org/10.1016/j.electstud.2023.102628

Scherpenzeel, A. C., & Das, M. (2010). 'True' longitudinal and probability-based internet panels: Evidence from The Netherlands. In M. Das, P. Ester, & L. Kaczmirek (Eds.), *Social and behavioral research and the Internet: Advances in applied methods and research strategies* (pp. 77–104). CRC Press, Taylor & Francis. https://www.lissdata.nl/

Somin, I. (2018). Knowledge about ignorance: New directions in the study of political information. In *The nature of belief systems reconsidered* (pp. 301–324). Routledge.

Tichelbaecker, T., Gidron, N., Horne, W., & Adams, J. (2023). What do we measure when we measure affective polarization across countries? *Public Opinion Quarterly*, *87*(3), 803–815. https://doi.org/10.1093/poq/nfad033

Torcal, M., & Harteveld, E. (2025). *Handbook of affective polarization*. Edward Elgar Publishing Limited. https://www.e-elgar.com/shop/gbp/handbook-of-affective-polarization-9781035310593.html#:~:text=This%20insightful%20Handbook%20provides%20a%20comprehensive%20overview%20of,issues%20of%20conceptualization%2C%20methodology%2C%20causes%2C%20consequences%2C%20and%20interventions

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, *69*, 102199. https://doi.org/10.1016/j.electstud.2020.102199

Wagner, M. (2024). Affective polarization in Europe. *European Political Science Review*, *16*(3), 378–392. https://doi.org/10.1017/S1755773923000383

Yeung, E. S., & Quek, K. (2025). Self-reported political ideology. *Political Science Research and Methods*, *13*(2), 412–433. https://doi.org/10.1017/psrm.2024.2

Address for correspondence: Jakob Kasper, Amsterdam School of Communication Research, University of Amsterdam, 1018 WV Amsterdam, The Netherlands.
Email: j.kasper@uva.nl

© 2025 The Author(s). *European Journal of Political Research* published by John Wiley & Sons Ltd on behalf of European Consortium for Political Research.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 15:53:12, subject to the Cambridge Core terms of use.