



## UvA-DARE (Digital Academic Repository)

### Public opinion on Twitter? How vote choice and arguments on Twitter comply with patterns in survey data, evidence from the 2016 Ukraine referendum in the Netherlands

van Klingeren, M.; Trilling, D.; Möller, J.

**DOI**

[10.1057/s41269-020-00160-w](https://doi.org/10.1057/s41269-020-00160-w)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

Acta Politica

**License**

Article 25fa Dutch Copyright Act

[Link to publication](#)

**Citation for published version (APA):**

van Klingeren, M., Trilling, D., & Möller, J. (2021). Public opinion on Twitter? How vote choice and arguments on Twitter comply with patterns in survey data, evidence from the 2016 Ukraine referendum in the Netherlands. *Acta Politica*, 56(3), 436-455.  
<https://doi.org/10.1057/s41269-020-00160-w>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.  
UvA-DARE is a service provided by the Library of the University of Amsterdam (<https://dare.uva.nl>)



# Public opinion on Twitter? How vote choice and arguments on Twitter comply with patterns in survey data, evidence from the 2016 Ukraine referendum in the Netherlands

Marijn van Klingeren<sup>1</sup> · Damian Trilling<sup>2</sup> · Judith Möller<sup>2</sup>

Published online: 30 April 2020  
© Springer Nature Limited 2020

## Abstract

Extensive research has been done on how social media have changed democratic society, politics, and public opinion. Social media are often regarded as a mirror of the public that, during political events, provides journalists and academics with a clear image of what position the public has on political issues and which sub-issues it uses to back it up. Yet, there is strong empirical evidence that active Twitter users differ in terms of background characteristics from the electorate, and that the most influential users possess specific traits. However, this does not necessarily mean that the opinions expressed on Twitter cannot reflect public opinion. This study aims to compare sub-issues used on Twitter to polled public opinion data in the context of the 2016 so-called 'Ukraine referendum' in the Netherlands. Our main findings indicate that there is a remarkable resemblance between the two domains in terms of sub-issues used and prominence of these sub-issues. Yet, this is mostly the case when not taking duplicates or retweets into account. Overall, the Twitter debate showed to be less nuanced than the polled public opinion data, as fewer sub-issues appeared.

**Keywords** Twitter · Social media · Public opinion · Political campaigns · Democracy

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1057/s41269-020-00160-w>) contains supplementary material, which is available to authorized users.

---

✉ Marijn van Klingeren  
[m.vanklingeren@ru.nl](mailto:m.vanklingeren@ru.nl)

<sup>1</sup> Sociology, Radboud Social Cultural Research (RSCR)/Interuniversity Center for Social Science Theory and Methodology (ICS), Radboud University Nijmegen, Postbus 9104, 6500 HE Nijmegen, The Netherlands

<sup>2</sup> Department of Communication Science, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands



Social media platforms have taken an important position in society, by forming a space of political and societal debate. It has been argued that they are one of the places where public opinion is formed and can be observed (Gayo-Avello 2013). Specific affordances of social media platforms, such as their near universal accessibility and open nature, create a low-threshold- and interactive communicative space. Twitter, in particular, has been the subject of extensive research on how new, social media have changed democratic society, politics, and public opinion (see Jungherr 2016 for an overview; Boyd 2010; Downey and Fenton 2003; Fuchs 2017; Papacharissi 2003, 2010). One important enigma in this strand of research is that, despite the fact that studies have repeatedly shown politically active Twitter users not to be representative of the general population (e.g., Barbera and Rivero 2015; Bode and Dalrymple 2014; McGregor 2019), and that social media are highly influenced by political players (Wells et al. 2016) academics and journalists rely on Twitter as a source of public opinion (Anstead and O’Loughlin 2015; Gayo-Avello 2013; Kreiss 2016). The rationale behind this choice is that the opinions and sub-issues represented on Twitter still provide insight into what everyday citizens think. But is the use of Twitter for this purpose really appropriate? In this study, we aim to address this question by comparing polled public opinion data on decisive reasons for the electoral decisions with the sub-issues expressed on Twitter.

Ideally, social media could provide an open forum for discussion in which the public can regain its democratic position, be critical and directly negotiate about decisions with other Twitter users, as well as journalists and politicians (e.g., Boyd 2010; Papacharissi and de Fatima Oliveira 2012). However, in reality, this open forum is only frequented by a selected group of users, of which only a small share actually partakes in political debates. Extant research suggests that these politically vocal users are quite different from those who abstain from Twitter or use it for different purposes (Jungherr 2016). They possess particular traits such as being more extravert and higher educated and more often male (Fuchs 2017). Hence, Twitter can by no means be regarded as a representative sample of the population from which public opinion can be simply derived. Research has engaged in understanding these audiences during political events, during which the active Twitter population was found to differ substantially from the entire electorate in terms of background characteristics (Bode and Dalrymple 2014; Jacobs et al. 2016; Vaccari et al. 2015).

Yet even though Twitter users are not representative of the population, their process of public opinion formation might still be reflective of the general process of public opinion formation, in particular as the opinions expressed on Twitter are in itself a part of public opinion formation. We therefore ask *whether Twitter makes as good of a proxy of public opinion during political events as polled public opinion data does*; the latter being the more established way to measure public opinion (Berinsky 2017).

We study this issue during a specific political event: the Dutch referendum on the Ukraine association treaty (“Ukraine referendum”). Public opinion in this case means sentiment in the form of vote-preference as well as the main sub-issues that are mentioned in relation to the referendum issue. This single-issue political event is a good case to compare variation in the use-, as well as in the prominence of certain referendum issue-related sub-issues, and so we aim to establish the extent to which



sub-issues on Twitter represent the same sub-issues as representative survey data do during the Ukraine referendum. To do so, we conducted a content analysis of tweets about the referendum, as well as a content analysis of open-ended survey questions from a representative sample of the population, which was conducted right after the events. Both are then systematically compared to investigate their overlap and divergences.

## Twitter, an ideal forum for public opinion formation?

Social media—like Twitter—have the potential to afford the ideal public sphere, in the way originally proposed by Habermas (1989, see also Anstead and O’Loughlin 2015 and Dahlberg 2011). He describes the public sphere [Öffentlichkeit] as an open forum for debate, to which all citizens have access. There is freedom of (unrestricted) expression about all relevant rules governing relations, and it is *the place where public opinion is formed*. In later work, Habermas (2006) nuanced this conceptualization, accounting for the critical role of mass-mediated communication. He explains “Moreover, the dynamics of mass communication are driven by the power of the media to select, and shape the presentation of, messages and by the strategic use of political and social power to influence the agendas as well as the triggering and framing of public issues” (Habermas 2006, p. 415). “As far as input from the outside is concerned, politicians and political parties are, of course, by far the most important suppliers... It follows that compared with politicians and lobbyists, the actors of civil society are in the weakest position.” (ibid. p. 419). The reality of social media, in particular the platform Twitter, is a manifest of this more nuanced conceptualization of the public sphere. The platform is open to all, yet influential political actors supply the most influential and visible content as a strategy to influence public opinion (Wells et al. 2016). Thereby the platform itself becomes a part of the complicated inter-media agenda setting process that underlies public opinion formation.

Nevertheless, social media facilitate communication in which every user can exchange information with different audiences (Gillmor 2004). Both public expression of one’s opinion as well as the interaction between individuals is represented and can be observed on social media. Interactive features provided by these media, such as the possibility to share or comment on specific content in varying degrees of “publicness” create a much more inclusive opportunity for the formation of public opinion (Stockmann and Luo 2017).

Fuchs (2017) argues, however, that describing social media as a platform of freedom and endless possibilities is too optimistic. First of all, because in these descriptions the necessary condition of *co-presence* that allow for private political activities to become part of the public sphere is often neglected (Fuchs 2017). Democratic deliberation and the formation of public opinion can only take place if *all can*, and *many do* participate in the debate. Despite its open nature, this is not always the case on Twitter. A growing body of studies (for an overview see Jungherr 2016) coherently shows that only small segments of the population take part in political debates.



In addition, universal access cannot be achieved, as the public sphere is only accessible to those who have the intellectual skills to partake on Twitter. For this reason, the sphere we find on Twitter can still be elite-driven. Moreover, people can block one another, therefore restricting access to part of the discussion for some. And, relying on social media as a source of public opinion may be problematic if bots or trolls interfere with the natural flow of the medium with strategic purposes; for instance, by liking and retweeting certain tweets, they may make some sub-issues seem more important than they actually are to their users.

## Twitter, a non-representative forum

In our particular case of the Netherlands, the Internet penetration rates are highest in the world. About 96% of the people in the Netherlands have access to an Internet connection (Internet World Stats 2018a, b), which means that about 13.7 million of the 14.3 million residents above the age of 15 have access to it. Among them, about 2.9 million people have a Twitter account, yet only little over a million (about 6% of those who have Internet access) are active on a daily basis (Van der Veer 2018). These numbers become even smaller when looking at *politically active* Twitter users (Bakker 2013); and this appears to be a sub-group with particular traits. Users that are engaged in political communication on Twitter are more politically interested (Bode and Dalrymple 2014; Vaccari et al. 2015), and/or are more often partisan (Barbera and Rivero 2015). In addition, they are younger, and males and students are overrepresented on Twitter (Bode and Dalrymple 2014; Vaccari et al. 2015).

Evidence for this Twitter population bias has been found in relation to the Ukraine referendum as well. A recent report showed that only 31% of the Dutch population actually used (any) social media as a source of information about the referendum.<sup>1</sup> And even though within the former group there was no gender difference, it did show a large overrepresentation of younger people (Jacobs et al. 2016). This is in line with the idea that those who do use social media during election campaigns are not representative of the electorate as a whole. On top of that, social media platforms such as Twitter are known to be at least partly shaped by so-called influencers.

Influencers can be described as highly present social media users, who play a central role in the distribution and filtering of information on political campaigns (Jungherr et al. 2011). In the context of political debates, these “influencers” are politicians as well as lobbyists and activists. The same actors are also likely to influence the mass-mediated debate either through Twitter or in addition to it (see Wells et al. 2016).

In sum, it is uncertain whether opinions on Twitter can be used as a proxy of public opinion. It can be argued that polling data are also not a valid reflection of considered public opinion, yet there is strong consensus among political scientists and

---

<sup>1</sup> This figure corresponds closely with recent statistics from Newman et al. (2019) on political social media use in the Netherlands (33%). These percentages are likely to be an overrepresentation of the true percentage as Internet users are most probably overrepresented in these surveys.



beyond that surveys are an accurate tool to efficiently tap public sentiment (Berinsky 2017). We therefore investigate whether there is any resemblance between polling and twitter data in order to see whether Twitter has the potential to be as good an indication of public opinion around election time as polls are. And so besides giving us an idea of the accuracy of Twitter as a proxy of public opinion, the comparison of Twitter- with survey data shows us whether representativeness of the units of analysis is truly necessary for a representative image of the general sentiment of among the public.

## The Ukraine referendum

On April 6th, 2016, the first-ever referendum took place after the introduction of the new Dutch referendum law in July 2015. This law allowed for Dutch citizens to initiate a consultative referendum about laws and treaties after these had been adopted by the Dutch senate and House of Representatives. As the Dutch referendum law was abolished again in 2018, our study makes use of this unique window of opportunity to investigate Twitter during a highly single issue-focused political event. This issue was *the approval of the association agreement between the European Union and Ukraine*. The referendum under study was initiated by a committee consisting of a prominent critical news-related website (GeenStijl.nl), a related action group (GeenPeil), and a group that describes itself as being anti-EU and pro-Europe (Burgercomité EU).

These Eurosceptic organizations were looking for a suitable legislation for a referendum, as soon as the new referendum law came into effect (Jacobs 2018). The treaty itself was meant to intensify (mostly economic) cooperation between EU member states and the Ukraine. A political–military association treaty was signed on 21th March between the heads of EU member states and non-elected Western Ukraine leader Arseniiy Yatsenyuk in an EU summit in Brussels. But no substantial attention was given to this at the time (Kleinnijenhuis et al. 2018). The issue was considered symbolic rather than substantial, with the main aim to question the legitimacy of the European Union (Heck 2016). The discussion surrounding this event seldom related to the actual referendum issue, but often hovered around (distantly) related sub-issues. This makes it an interesting case to investigate more closely, as it is a political issue in which partisanship plays less of a role than in regular election campaigns. Yet, influential actors in the campaign were likely to also use Twitter as a communication mode to direct voters (see Well et al. 2016). Their posts triggered and organized the public debate, in particular those by the blog that initiated the referendum (GeenStijl), as well as set the overall agenda of the public debate.

On the day of the referendum, voters were asked the question “Are you in favor of or against the law that approved the Association Treaty between and the European Union and the Ukraine?” Turnout was low, yet above the 30% threshold. The lack of participation can be attributed to the non-binding yet complex nature of the referendum. The outcome of the referendum showed a clear ‘no,’ with 61% voting against and 38.21% voting in favor of the treaty and 0.79% casting a blank vote. To



see whether Twitter represents public opinion in terms of the referendum outcome, we ask the following research question:

**RQ1** To what extent are proponents or opponents over- or underrepresented on Twitter compared to the actual referendum outcome?

Moreover, several issues were stresses during the campaign leading up to the referendum. These included the sub-issue that this treaty was simply about trade, or that it served as a prelude to EU accession of the Ukraine, or that the treaty would ‘provoke’ Putin and that the Ukraine was a corrupt country (Jacobs 2018; Kleinnijenhuis et al. 2018). We further explore our main thesis by addressing the use of these sub-issues in the following research question:

**RQ2a** Do the sub-issues mentioned on Twitter in relation to the Ukraine referendum in the Netherlands differ or not from the sub-issues in the polling data?

Moreover, if there is overlap in the sub-issues, we wonder if these are similar in their level of prominence. In other words, if Twitter users pay just as much attention to these sub-issues as the people in the polls did.

**RQ2b** Do the sub-issues mentioned on Twitter differ in their *prominence* from those mentioned in the polling data?

Although understanding the prominence and overlap between the sub-issues overall will provide us with important insights, Kleinnijenhuis, van Attenveldt, and Dekkers (2018) argue that the pro- and con-campaign highlighted their own specific sub-issues. This depended on whether they could use these sub-issues to make an argument in the direction of their stance. It also relates to the notion of issue ownership, a line of literature which states that voters tend to vote for parties that they find competent to handle certain issues (Green and Hobolt 2008; van der Brug 2004). Hence, these parties will emphasize the issues that direct the voter to their party. Even though in this paper our interest is not in party position, emphasis on particular sub-issue might just as well navigate voters in a particular direction. Meaning that the overlap in sub-issues between the two domains become better visible if we split up the pro- and con-voters/ tweets.

**RQ3a** Do the sub-issues mentioned by the pro/con-camp on Twitter in relation to the Dutch Ukraine referendum differ or not from the sub-issues mentioned by the pro/con camp in the polling data?

**RQ3b** Do the sub-issue mentioned by the pro/con-camp on Twitter differ in their *prominence* from those mentioned by the pro/con-camp in the polling data?

Finally, we are interested in the influence of interactivity and peer-curation in the process of public opinion formation on Twitter. Posts can gain visibility beyond



the network of a single user if it is shared (retweeted) by those who initially saw the posts, especially since they are often part of larger communication strategy that spans beyond Twitter. In this way a single post can gain a much larger audience. In addition, being shared is an important cue in determining the quality of a post (Kupavskii et al. 2012). That means that a Tweet that is shared many times not only has more visibility but is also more convincing. Moreover, social media messages that are shared more often, also tend to hold higher news value, implying that they are more influential even without the popularity cues (Trilling et al. 2017) Having said that, while these mechanisms are important in the *formation* of public opinion, it is arguable whether including sharing dynamic improve the *observation* of public opinion, as they decrease the relative importance of voices with a smaller following, we therefore ask:

**RQ3c** Does including shared tweets improve the accuracy of the observation of public opinion on Twitter?

## Data and method

The data for this paper were gathered around the time of the referendum that took place on April 6, 2016, and consists of both public opinion and Twitter data. The former were gathered the day of, up to 20 days after the referendum took place. The Twitter data were gathered a week before, until the day of the referendum.

### Data collection survey

Online data collection took place with the use of LISS panel (Longitudinal Internet Studies for the Social sciences) administered by CentERdata. This is a representative database of the Dutch adult population, which includes over 4500 households comprising about 7000 people who asked to participate in monthly surveys. Invitation to the survey occurred via e-mail. To reduce noncompliance, financial incentives were given after completion. To achieve a representative pool of participants in the LISS panel, they apply non-probability sampling and provide people with devices to fill out questionnaires in case they are not in the possession of an appropriate device. For our survey, a random sample was drawn from this database of unique household members. A total of 1523 respondents were invited to participate of which 1262 responded to the request and 1222 completed the questionnaire (a total response rate of 80.2%). Eventually, 1168 answered the questions with regard to the Ukraine referendum (item non-response of 4.6%).

### Representativeness

Given the fact that we use polling data as a baseline, it is important that these data are in fact representative of the electorate. Overall, Table B1 (Appendix B in Electronic supplementary material) shows the descriptive statistics of our data compared





to the entire Dutch population. This shows that after attrition, there is a slight overrepresentation of people age 40 to 80, and higher educated, while those with a mid-range educational level are underrepresented. Generally, voters tend to be slightly older and slightly higher educated than the average population (Goerres 2007; Hillygus 2005). However, to be certain, we ran the analyses while weighing the data for age. This yielded the same results (see Appendix C, Figures C1 and C2 in Electronic supplementary material).

## Variables

Two survey questions were used for the analyses. The first one concerns vote choice: “on April 6, a referendum was held on the association treaty between the EU and Ukraine, what did you vote?” 1 ‘pro-treaty,’ 2 ‘con-treaty,’ 3 ‘a blank vote,’ 4 ‘I did not vote,’ 5 ‘I was not allowed to vote,’ 6 ‘I’d rather not say.’ In our survey, 55% indicated that they voted. This means voters are overrepresented in our sample, or respondents overreported their electoral participation due to social desirability. Out of those who voted, 42.5% voted in favor of the treaty, and 56.0% voted against the referendum. These shares are close to the official result, but our sample is biased by around 4 to 5 percentage points.<sup>2</sup>

The second open-ended question measures the sub-issues discussed on Twitter and polling data, after the respondent’s vote choice was asked, is “What was the most important reason for you to abstain from voting/vote for/vote against/vote blank?”. This variable was coded similarly as the Twitter data and is explained in more detail below.

## Data collection Twitter

We collected our Twitter sample by using the Twitter streaming API and selecting all tweets in which at least one of the following terms was included (English translations between brackets) Oekraïne (‘Ukraine’) and the hashtags ‘#referendum,’ ‘#oekraïnerferendum,’ ‘#geenpeil’ (the main initiator), ‘#stemvoor’ (‘votepro’), ‘#stemtegen’ (‘votecon’), ‘#stemja’ (‘voteyes’), ‘#stemnee’ (‘voteno’), and #associatieverdrag (‘#association treaty’) between March 28 and April 2nd, 2016. In total, we collected 30,828 tweets. After collapsing all retweets and creating a variable that indicated how often retweets of the tweet occurred in the dataset, we were left with 16,120 unique tweets. Out of these, we drew a random sample of 3000 tweets. 780 Of them (26%) were found irrelevant, as they mostly concerned another, simultaneous referendum.

---

<sup>2</sup> To check whether this affected our results, we ran our analysis on a weighted dataset to correct for the overrepresentation of this sub-group. The results remained intact and highly compatible.



## Coding

The relevant tweets ( $n = 2220$ ), as well as the open-ended question ( $n = 1168$ ), were systematically hand-coded by the three researchers. The hand-coding method was chosen because of the implicit nature of some of the mentioned sub-issues. This is something that is difficult to extract automatically. For this purpose, we applied an inductive approach, which entails that we started by narrowing down some of the sub-issues that were considered relevant in the debate on the Ukraine referendum from our Twitter and survey data separately. These sub-issues were then written down in two codebooks (see Appendix A in Electronic supplementary material), which was supplemented with additional sub-issues during the first coding stage in the case that the aforementioned codebook did not suffice.

Subsequently, the coding process restarted until all relevant sub-issues were coded adequately. In-between each step, the researchers came together to discuss the additional sub-issues and define the categories. This method, as well as the fact that, in contrast to the Twitter data, the survey initially also contained information from people who did not vote or voted blank, led to a much more elaborate coding scheme for the survey- than for the Twitter data. The eventual list showed seven sub-issues with overlap between the two domains: (1) the European Union, (2) Russia, (3) the Dutch government, or governing politicians, (4) the Ukraine people, (5) corruption, (6) immigration/visa sub-issue, and (7) sub-issues related to the financial costs/benefits for the Netherlands and or Europe. Moreover, sub-issues that only prevailed on Twitter in relation to the referendum concerned (1) Ukraine Nazis, (2) civil war, (3) and turnout of the referendum. The survey data in addition showed sub-issues related to (4) collaboration, (5) reasons why people could not vote/low interest (6) no clear picture on sub-issue of referendum, (7) little faith in situation, referendum or the like, (8) the strategic vote, and (9) low efficacy, (10) against referendum, (11) representatives should decide for them. For the analyses, these data were aggregated in order to allow a comparison between their relative presence in the polling and in Twitter data.

## Inter-coder reliability

Of the polling data the inter-coder reliability test was performed on a random sub-sample of 118 questions (10% of the total data), which led to an average Krippendorff's alpha of 0.85. Of all sub-issues, one was below the threshold of 0.68.<sup>3</sup> Of the Twitter data, a random sub-sample of 172 tweets (7.8% of the coded Twitter data) was selected for the inter-coder-reliability test, which averaged at 0.72. Two low-prominence sub-issues were excluded from the analysis because of their consistently

<sup>3</sup> The lowest one of .5 relates to the 'strategic vote' sub-issue, which was categorized only twice. In both cases, a different duo of coders categorized it as such, which indicates that this was not a case of disagreement, but of a missed identification. For this reason and because it is not one of the overlapping themes (this category will not be used in most of the analyses), we think that the low inter-coder reliability is not problematic.



low score on the reliability test throughout the process.<sup>4</sup> One category (sub-issues related to the Dutch government) was kept despite its low reliability score, because of its prominence in the data (Krippendorff's alpha of 0.43). In a debriefing session, one of the coders noticed the systematic error in their approach to this frame. After reaching agreement over the content of this sub-issue, the category became so clear-cut that we addressed this problem by writing a brief script<sup>5</sup> to identify the sub-issues automatically.

## Results

In this paper, we posed the question to what extent Twitter is a good proxy of public opinion with regard to the outcome of the Ukraine referendum. Research question 1 asked: To what extent are proponents or opponents over- or underrepresented on Twitter compared to the actual referendum outcome? The electoral result was 38.2% in favor and 61.0% against the treaty with Ukraine (Kiesraad 2016), with a turnout of 32.28%. When we look at the tweets that explicitly call for voting for or against the treaty, the error is around 6–7 percentage points—31.1% could be classified as 'pro,' 67.3% as against (based on unique tweets)—but still close to the actual outcome. If we include retweets and tweets with duplicate text, the figures are virtually the same (31.9% vs 66.4%). Thus, we can answer our first research question: Regarding the ratio of tweets advocating voting for versus voting against the treaty, we see a slight bias in favor of the 'against' camp, but overall there is a resemblance. Yet, our polling data was even slightly closer to the actual outcome: with 42.5% pro and 56.0% against, the error is around 5 percentage points. Despite the fact that the survey grossly overrepresented voters (55.3%) compared to non-voters, which is 71% more than the actual turnout, Twitter is slightly less reliable in terms of outcome.

Research question 2a asked: Do the sub-issues mentioned on Twitter in relation to the Ukraine referendum in the Netherlands differ or not from the sub-issues in the polling data? In general, it seems that there is a considerable overlap in the use of sub-issues (see Fig. 1 and Table 1)—with a couple of exceptions. On Twitter, there was a largely retweeted tweet about Nazi tendencies in Ukraine, which was not mentioned in our survey responses. Vice versa, we observed several categories such as strategic voting, the Ukrainian civil war, testimonials, and international cooperation in (few) survey responses, but not in our Twitter dataset. In sum, in terms of the sub-issues mentioned, there is still a vast amount of overlap.

Moving onto research question 2b: do the sub-issues mentioned on Twitter differ in their prominence from those mentioned in the polling data? Comparing the prominence of the sub-issues that are both discussed on Twitter and in the polling data<sup>6</sup>,

<sup>4</sup> The excluded categories are "sub-issues related to the Ukrainian civil war," and "sub-issue related to turnout of the referendum."

<sup>5</sup> All scripts are available at [URL MASKED FOR BLIND PEER REVIEW].

<sup>6</sup> In the survey we asked voters whether they voted in favor- or against the treaty. But we also asked people why they did not vote or vote blanc, which naturally means the range of topics is much broader in the survey than the Twitter data. An overview of the relevance of all topics in the survey- compared to the Twitter data can be viewed in Figure B1 of appendix B in Electronic supplementary material.



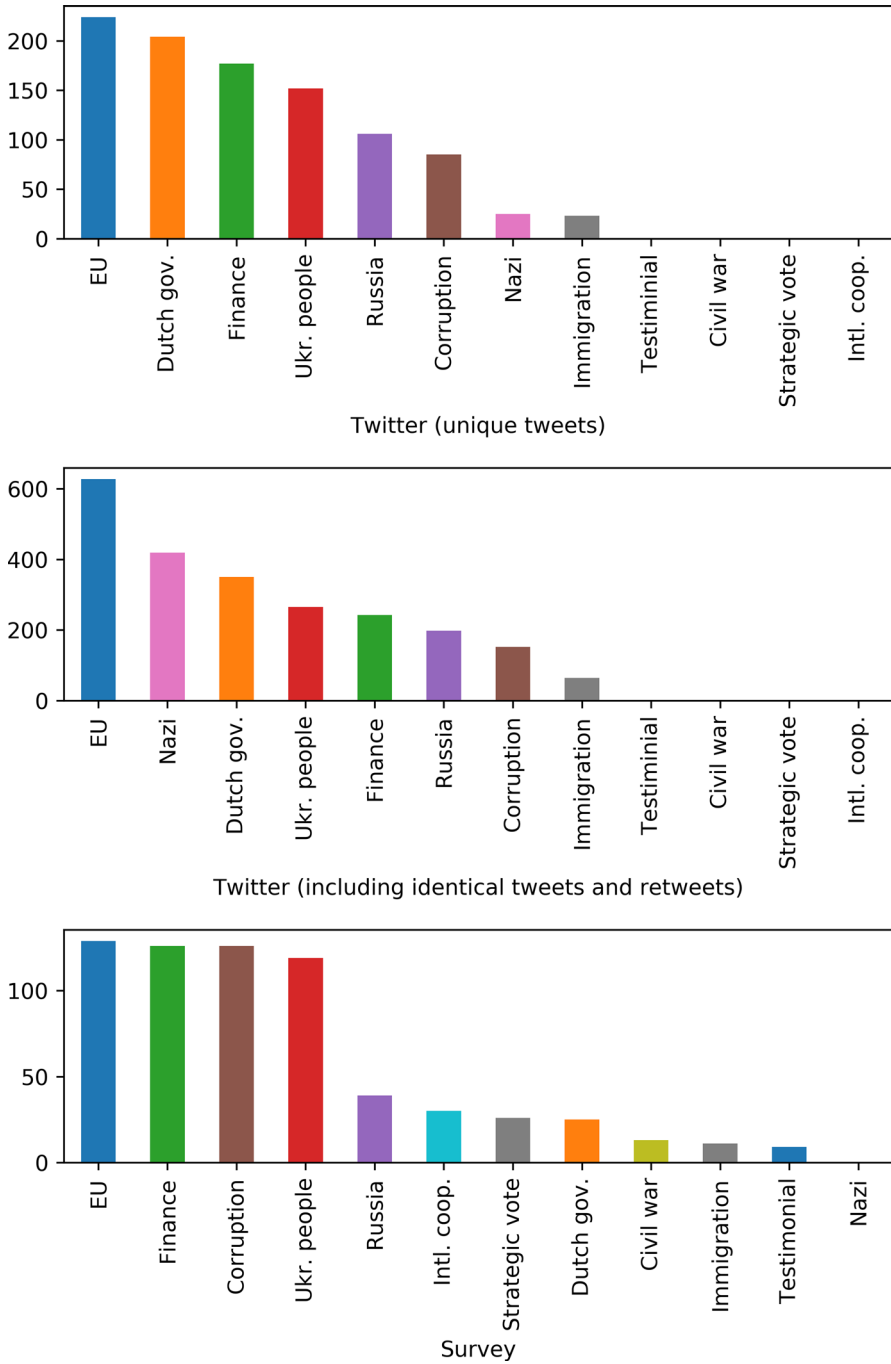


Fig. 1 Absolute presence of sub-issues in Twitter and polling data, including sub-issues that do not overlap between the two data sources



**Table 1** Relative presence of sub-issues in Twitter and polling data, including sub-issues that do not overlap between the two data sources

	Twitter [(unique %)]	Twitter [(including identical tweets and retweets) %]	Polling data (%)
EU	24.6	27.1	19.8
Dutch government	22.4	15.1	3.8
Finance	19.4	10.4	19.3
Ukrainian People	16.7	11.4	18.2
Russia	11.6	8.5	6.0
Corruption	0.0	6.6	19.3
Nazi	2.7	18.1	0.0
Immigration	2.5	2.8	1.7
Testimonial	0.0	0.0	1.4
Civial war	0.0	0.0	2.0
Strategic vote	0.0	0.0	4.0
International cooperation	0.0	0.0	4.6

we see that there is some remarkably compatibility between the two domains (see Fig. 2). At first glance, we see that the most prominent issues are relatively similar (see also Table 2). Whereas the distribution in the Twitter dataset shows more continuity regarding their frequency than the survey data, this difference should not be overstated. For the data underlying Fig. 2, we calculated the Herfindahl–Hirschman Index. Higher values indicate a stronger concentration (i.e., few issues gaining most of the attention), but the index of the two domains is remarkably similar and shows little domination of particular topics (HHI polling data = 0.20, HHI unique tweets = 0.17, HHI twitter = 0.17).

Framing the main issue as an EU issue seems to be dominant in both domains. Also, discussing it in terms of the consequences for the Ukrainian people seems to be of comparatively high importance. The relationship with Russia is present, but a rather minor sub-issue in both domains, while immigration and visa-related sub-issues are generally considered to be of limited importance.

Next, to these commonalities, we also see a couple of remarkable differences. The role of the Dutch government was only of minor importance to the survey respondents but was often mentioned in the tweets. Reversely, corruption—very frequently mentioned in the surveys—was only a minor sub-issue on Twitter. If we correlate these agendas using Spearman's rho rank correlation coefficient, we can confirm what we already visually saw in Fig. 1. The unique Twitter agenda and the survey agenda are moderately correlated ( $r=0.49$ ), as well as duplicate tweets/retweets and the survey agenda ( $r=0.43$ ). The two Twitter agendas are highly positively correlated with one other ( $r=0.96$ ).

There was, however, one sub-issue that went viral on Twitter that was never mentioned in the survey (see Appendix B, Figure B2 and Table B2 in Electronic supplementary material). This related to Ukrainian Nazi's and shows that even though only a few tweets addressed the issue, they were retweeted very often, thereby distorting



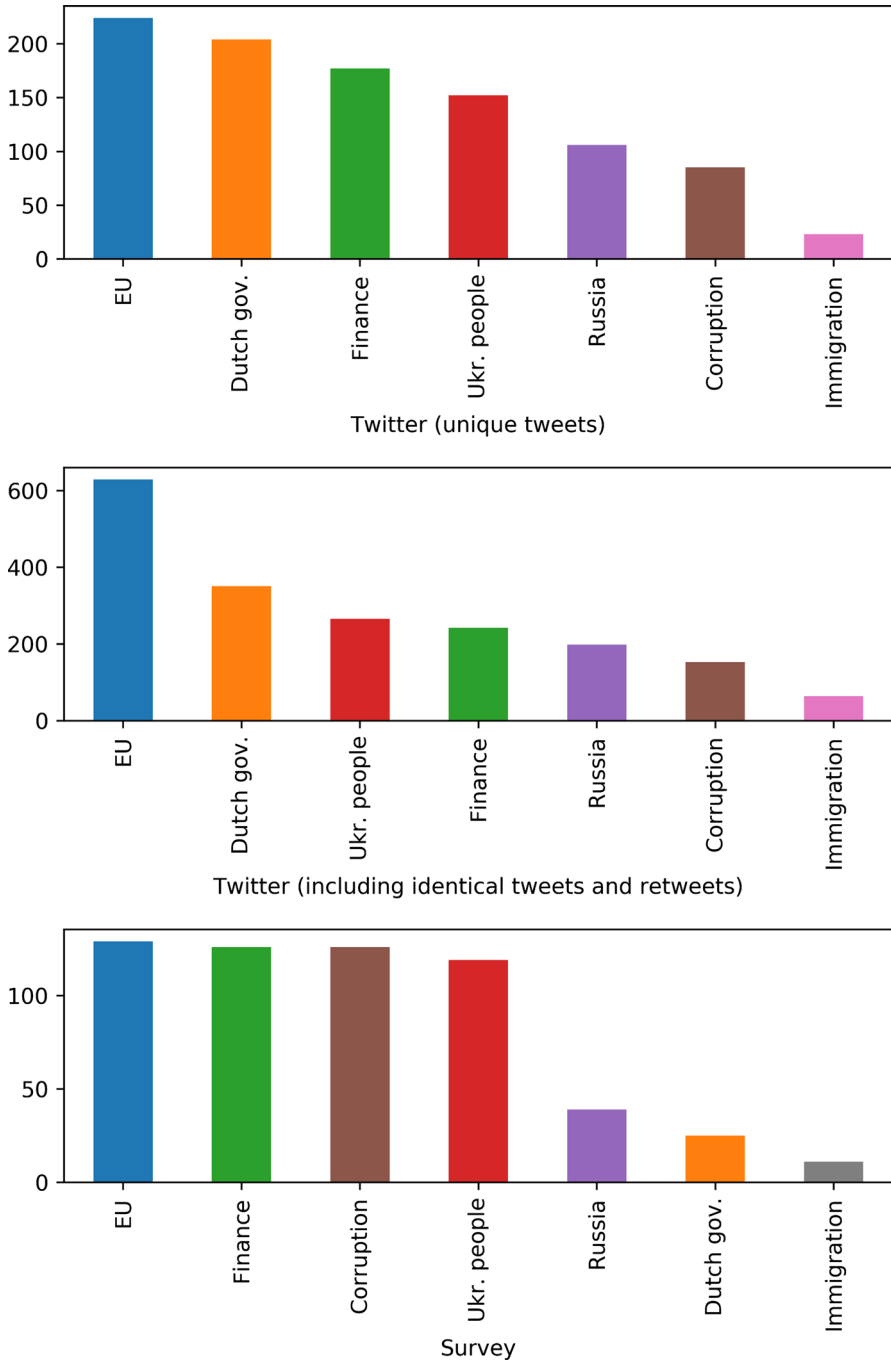


Fig. 2 Absolute presence of sub-issues in Twitter and polling data, excluding sub-issues that do not occur in both corpora



**Table 2** Relative number of tweets and polling responses mentioning of sub-issues, excluding survey-only sub-issues

	Twitter [(unique) %]	Twitter [(including identical tweets and retweets) %]	Survey (%)
EU	23.1	33.1	22.4
Dutch government	21.0	18.4	21.9
Finance	18.2	14.0	21.9
Ukrainian People	15.7	12.7	20.7
Russia	10.9	10.4	6.8
Corruption	8.8	8.0	4.3
Immigration	2.4	3.4	1.9

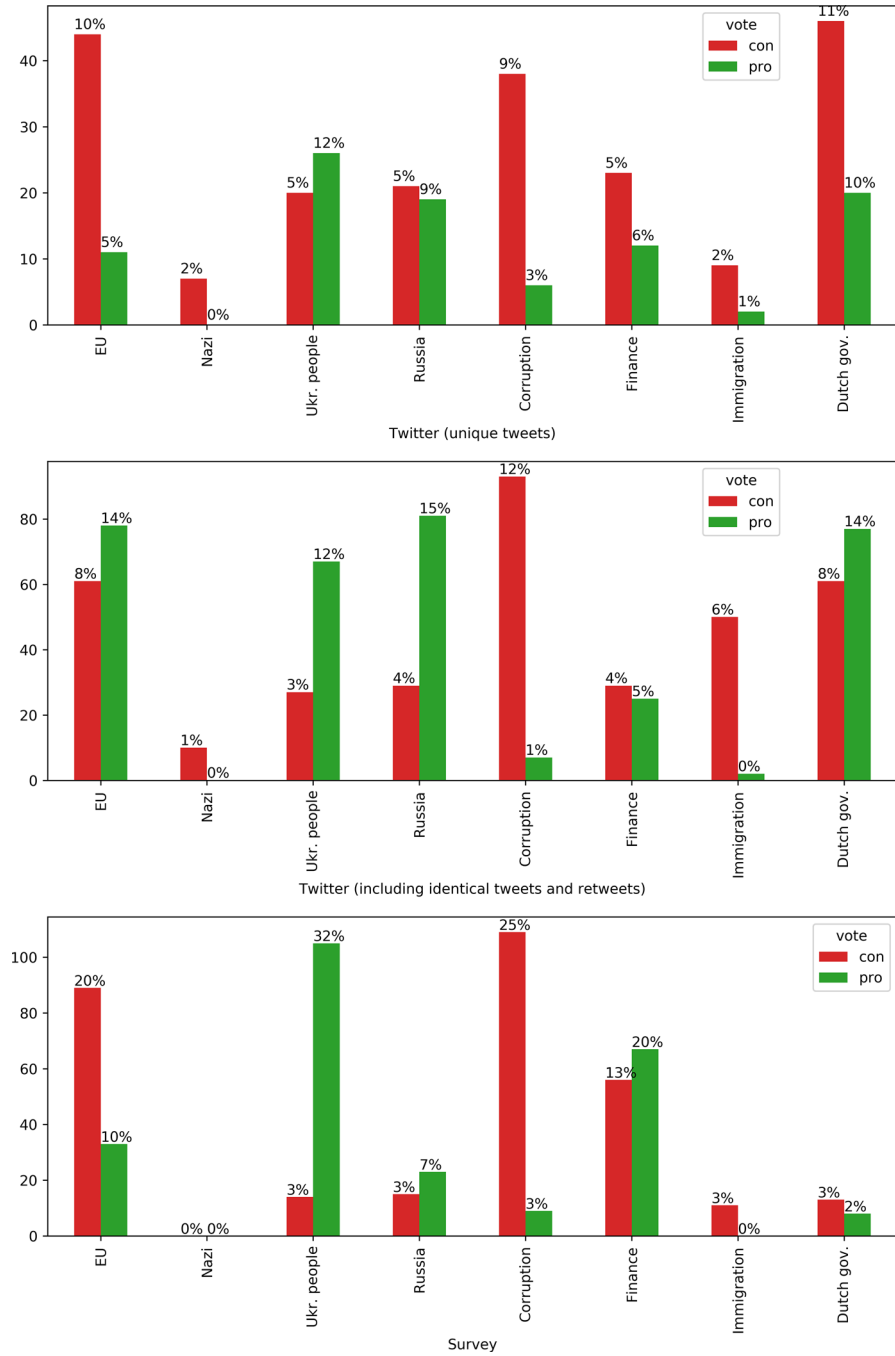
the overall picture on Twitter. If we look at the correlations when this ‘Nazi’-sub-issue is taken into the equation, which was one of the sub-issues that was left out initially because it did not show in the polling data, we see that the unique Twitter data suddenly correlates much more with the polling data ( $r=0.63$ ), as the rank order between the two domains becomes more aligned. While the correlation between the polling data with the duplicate Twitter data ( $r=0.12$ ) as well as of the the correlation between the polling data with the unique Twitter agenda inevitably goes down ( $r=0.62$ ). The implications of these differences will be discussed in more detail in the “[Discussion](#)” section. For now, there is some evidence for the idea that Twitter resembles the survey data to some extent.<sup>7</sup>

Turning to research question 3a—Do the sub-issues mentioned by the pro/con-camp on Twitter in relation to the Dutch Ukraine referendum differ or not from the sub-issues mentioned by the pro/con-camp in the polling data?— we see in Fig. 3 that the sub-issues overlap (there is none that is used exclusively by one side), but their magnitude differs considerably.<sup>8</sup> And although they tend to use the same arguments overall, immigration is not mentioned by the pro-camp in the polling data, where it is on Twitter. So, when moving onto answering research question 3b: Do the sub-issue mentioned by the pro/con-camp on Twitter differ in their *prominence* from those mentioned by the pro/con-camp in the polling data? We see some

<sup>7</sup> One may object to our analysis, given that our survey was slightly biased with regard to age and outcome. As an additional robustness check, we therefore carried weighing for these variables. If we weigh for age (effectively correcting for the overrepresentation of respondents between 50 and 70 in our sample), we find a very similar pattern as without (see Appendix C Figure C1 in Electronic supplementary material), leading to very few changes. We see that correcting for the age distribution does not change our conclusion. Similarly, correcting for the vote distribution does not change the general picture either (see Appendix C Figure C2 in Electronic supplementary material). While the absolute numbers of mentioned sub-issues change slightly, the relative importance does not. In any scenario, the topics Corruption, EU, and Finance form the top 3 in the survey data and are virtually indistinguishable.

<sup>8</sup> There are some more interesting observations to be made. For instance, while someone who followed the political developments closely may have expected that the role of former heroes of the Maidan revolution (as part of the Ukrainian people sub-issue), or the geopolitical stakes of Russia (as part of the Russia sub-issue) may have dominated the discourse, this does not seem to be the case.





**Fig. 3** The absolute and relative presence and overlap in sub-issue occurrence on Twitter and in survey data split by pro- and con-voters (bars resemble absolute numbers)





distinct patterns for each side, as there are two sub-issues that are used only by the 'con' camp (corruption and immigration), both on Twitter and in the survey. While any references to the Ukrainian people are almost exclusively used by supporters of the referendum in the survey data.

It is interesting that most other sub-issues are in fact used by both supporters and opponents. Regarding Twitter, it seems that the decision whether or not to eliminate retweets influences the conclusions quite a lot. For example, there are few unique pro-tweets about the EU, but these tweets appear to be retweeted often. This reflects in the correlation we find between Twitter and survey data on the pro-side. This is 0.68 for the unique tweets and 0.54 for those including retweets. For the con-side, these differences are much smaller, but again the unique tweets show the highest correlation 0.62 versus 0.59, respectively. In terms of diversity of the issue agenda when divided by pro- and con-camps, we observe more concentration (=less diversity) in the survey responses ( $HHI_{pro}=0.20$ ,  $HHI_{con}=0.25$ ) than in unique tweets ( $HHI_{pro}=0.19$ ,  $HHI_{con}=0.16$ ) and all tweets ( $HHI_{pro}=0.21$ ,  $HHI_{con}=0.16$ ). The graphical representation in Fig. 3 underlines that the concentration in the survey responses is not so much due to an extreme overrepresentation of the top categories, but mainly due to absence (Nazi) or nearly absence (immigration; Dutch government) of some sub-topics in the survey data.

Finally, RQ3c asked whether including shared tweets improves the accuracy of the observation of public opinion on Twitter. As we have seen, the unique Twitter data correlates higher with the survey data than the Twitter dataset that included retweets. In particular, as the stark difference between the frequency of the Nazi sub-issue illustrated, we saw that a very low number of tweets can, in case they go viral, bias the overall picture.

## Conclusion

Since their arrival social media platforms have made people wonder about the influences of these media forms on democratic society, politics, and public opinion (see Boyd 2010; Fuchs 2017; Papacharissi 2010). Ideally, social media platforms provide space for open deliberation and a sphere in which public opinion can form freely. This raises the empirical question whether the public debate on social media is a reliable source of public opinion. In this study, we addressed this issue to answer the question *whether Twitter makes as good of a proxy of public opinion during political events as polled public opinion data does*. This question is of particular relevance, since journalists and academics often rely on Twitter as a source of public opinion (Gayo-Avello 2013; Kreiss 2016; McGregor 2019). We addressed this question by comparing polled public opinion data to Twitter data during the Dutch Ukraine referendum (2016). Specifically, we focused on observing vote preference expressed on Twitter and the main sub-issues that are discussed in relation to this referendum.

Our findings can be distilled in four key insights: First, Both Twitter and the survey data diverged from the actual referendum results, yet the survey data were one percentage point closer to the actual outcome. Second, in the survey data, we



observe a wider range of sub-issues associated with the referendum by the respondents; yet, the sub-issues put forward by Twitter users and survey respondents resembled each other, also in terms of prominence. This overlap became somewhat stronger when we distinguished between the sub-issues of pro- and con-voters. And finally, unique tweets showed a consistently stronger match with our polling data than the Twitter data that included retweets.

We conclude that Twitter can provide us with an imperfect, but reasonably reliable proxy of public opinion. However, to achieve this, it is important to exclude retweets. Including those leads to situations in which case some extremely often shared tweets exert a disproportionate influence. For example, one specific tweet mentioning nazis was so widely shared, that its inclusion of retweets meant that the correlation between the polling data and the duplicate Twitter data became barely significant. Our analysis also revealed that the Twitter discussion was more focussed around a smaller number of sub-issues, meaning that if it is used as a proxy for public opinion, one loses part of the nuance of the debate. Ideally, one also needs to take partisanship/ vote choice into consideration during political events, which leads to higher congruence with the survey data.

## Discussion

Despite its clear contributions toward understanding the role of Twitter in the political domain, this paper faces some limitations which we discuss below. First of all, the data collection took place during a single and rather unique event. Although this provided us with important information, it raises questions about the generalizability of the results beyond this specific case. We believe that around political events of a different nature, such as national or local elections, we will see similar patterns, especially because we expect that the group of active Twitter users around such events is rather similar. The sub-issues will, of course, be context-dependent, but we see no reason to assume that the overlap between polling data and Twitter data will vary from event to event. Outside of election periods, however, we might find different results, but that would be subject to a whole new range of analytical problems.

Furthermore, by asking respondents to name sub-issues in an open question, we might have missed issues that were relevant to the respondents but that they could not remember as they filled in the survey. However, had we provided a list of issues, we might have guided the respondents too much. It is also important to note that we were forced to make a tradeoff regarding the scope of the sub-issues. A sub-issue like “Russia,” for instance, combines arguably even more distinct sub-issues like international relationships *with* Russia and the geopolitical stakes *for* Russia. Lumping together, these may have resulted in an over-estimation or overlap. On the other hand, further splitting up would have resulted in very small cell sizes that may not have been informative any more either.

Based on our findings, we can identify several avenues for further research. The first would be to use the same coding scheme to analyze the entire media discourse. One way in which social media effects can travel is via traditional news media. Analyzing these media in the same systematic way as we analyzed



Twitter data would allow to better understand the interplay between all three arenas (social media, news media, and public opinion as captured in survey data). It would be even more interesting to take into account cross-time and space dynamics. For instance, using a network approach, one could model cascades to understand how sub-issues spread both within and across the three arenas.

As we have seen, a single tweet can—if going viral—have a noticeable impact on a social media dataset. In general, it seems that the number of unique tweets in our dataset compared to the number of retweets was relatively low; however, none of this tells us much about the *actual* reach of the tweets. After all, we know nothing about how many people have been exposed to, let alone processed any given tweet. Yet, this would be necessary for a more comprehensive understanding of the exact contribution of a given tweet to the public sphere. Somewhat relatedly, it is also a contested issue how many tweets are actually written by trolls or bots. In the end, we would argue, this is irrelevant when we want to study how opinion on Twitter differs from opinion as expressed in survey data; nevertheless, it is of crucial importance if we want to understand the dynamics of public opinion formation.

Despite its restrictions, our findings provide vital new insights with regard to the role of Twitter in the political process and its role in democracy. It shows us how the two data domains are related to one another and in the end is a robustness check of survey data as well as Twitter data. As both are non-perfect ways to get a glimpse of public opinion. Limitations aside, we conclude that Twitter shows an accurate and a distorted mirror of public opinion. While the sub-issues used are comparable, a single viral tweet can distort the image, and in addition the discussion on Twitter is more focus and therefore may lack some of the nuance which is important to understand what the public thinks of as important during political events.

## Compliance with ethical standards

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

## References

- Anstead, Nick, and Ben O'Loughlin. 2015. Social Media Analysis and Public Opinion: The 2010 UK General Election. *Journal of Computer-Mediated Communication* 20 (2): 204–220.
- Bakker, Tom. 2013. *The myth of the active online audience?* (Doctoral dissertation).
- Barbera, Pablo, and Gonzalo Rivero. 2015. Understanding the Political Representativeness of Twitter Users. *Social Science Computer Review* 33: 712–729.
- Berinsky, A.J. 2017. Measuring public opinion with surveys. *Annual Review of Political Science* 20: 309–329.
- Bode, Leticia, and Kajsa E. Dalrymple. 2014. Politics in 140 characters or less: Campaign communication, network interaction, and political participation on Twitter. *Journal of Political Marketing*. <https://doi.org/10.1080/15377857.2014.959686>.



- Boyd, Danah. 2010. Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In *Networked Self: Identity, Community, and Culture on Social Network Sites* (ed. Zizi Papacharissi), pp. 39–58.
- Cameron, Jaelyn, and Nick Geidner. 2014. Something Old, Something New, Something Borrowed from Something Blue: Experiments on Dual viewing TV and Twitter. *Journal of Broadcasting and Electronic Media* 58 (3): 400–419.
- Dahlberg, Lincoln. 2011. Re-constructing digital democracy: An outline of four 'positions'. *New Media & Society* 13: 855–872.
- Downey, John, and Natalie Fenton. 2003. New Media, Counter Publicity and the Public Sphere. *New Media and Society* 5: 185–202.
- Enli, Gunn. 2017. New Media and Politics. *Annals of the International Communication Association* 41 (3–4): 220–227. <https://doi.org/10.1080/23808985.2017.1392251>.
- Fuchs, Christian. 2017. *Social Media: A Critical Introduction*. New York: SAGE Publications Ltd.
- Gayo-Avello, Daniel. 2013. A Meta-Analysis of State-of-the-Art Electoral Prediction From Twitter Data. *Social Science Computer Review* 31: 649–679.
- Gillmor, Dan. 2004. *We the Media: Grassroots Journalism by the People, for the People*. Sebastopol, CA: O'Reilly Media Inc.
- Goerres, Achim. 2007. Why are Older People More Likely to Vote? The Impact of Ageing on Electoral Turnout in Europe. *The British Journal of Politics and International Relations* 9: 90–121.
- Green, J., and S.B. Hobolt. 2008. Owing the Issue Agenda: Party Strategies and Vote Choices in British Elections. *Electoral Studies* 27 (3): 460–476.
- Habermas, Jurgen. 1989. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Cambridge: Polity.
- Habermas, J. 2006. Political Communication in Media Society: Does Democracy Still Enjoy an Epistemic Dimension? The Impact of Normative Theory on Empirical Research. *Communication Theory* 16: 411–426. <https://doi.org/10.1111/j.1468-2885.2006.00280.x>.
- Heck, W. 2016. Oekraïne kan ons niets schelen. In *NRC*. <https://www.nrc.nl/nieuws/2016/03/31/oekraïne-kan-ons-niets-schelen-1606419-a969298>. Accessed 10 July 2018.
- Hillygus, D.Sunshine. 2005. The Missing Link: Exploring the Relationship Between Higher Education and Political Engagement. *Political Behavior* 27: 25–47.
- Hindman, Matthew. 2008. *The Myth of Digital Democracy*. Princeton: Princeton University Press.
- Internet World Stats. 2018a. *Top 50 Countries with the Highest Internet Penetration Rates*. Retrieved from, <https://www.internetworldstats.com/top25.htm>
- Internet World Stats. 2018b. *Internet Stats and Facebook Usage in Europe December 2017 Statistics*. Retrieved from, <https://www.internetworldstats.com/stats4.htm>
- Jacobs, Kristof. 2018. Referendums in Times of Discontent. *Acta Politica* 53: 489–495.
- Jacobs, Kristof, Marijn van Klingeren, Henk van der Kolk, Tom van der Meer, and Eefje Steenvoorden. 2016. *Het Oekraïne-Referendum: Nationaal Referendum Onderzoek 2016*. Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.
- Jungherr, Andreas. 2016. Twitter Use in Election Campaigns: A Systematic Literature Review. *Journal of Information Technology and Politics* 13 (1): 72–91.
- Jungherr, Andreas, Pascal Jürgens, and Harald Schoen. 2011. Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., and Welpe, I. M. "Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment". *Social Science Computer Review* 30: 229–234.
- Kiesraad. 2016. Uitslag Referendum Associatieovereenkomst met Oekraïne Onherroepelijk. <https://www.kiesraad.nl/actueel/nieuws/2016/04/19/uitslag-referendum-associatieovereenkomst-met-oekraïne-onherroepelijk>
- Kleinnijenhuis, Jan, and Wouter Van Atteveldt. 2016. The Impact of the Explosion of EU News on Voter Choice in the 2014 EU Elections. *Politics and Governance* 4 (1): 104–115. <https://doi.org/10.17645/pag.v4i1.471>.
- Kleinnijenhuis, J., W. van Atteveldt, and V. Dekkers. 2018. Partial Priming: How Issue News Shapes Issue Saliency, Which Shapes Turnout but not the Vote. *Acta Politica* 53 (4): 569–589.
- Kreiss, Daniel. 2016. Seizing the Moment: The Presidential Campaigns' Use of Twitter During the 2012 Electoral Cycle. *New Media and Society* 18 (8): 1473–1490.
- Kupavskii, A., Ostroumova, L., Umnov, A., Usachev, S., Serdyukov, P., Gusev, G., & Kustarev, A. (2012, October). Prediction of Retweet Cascade Size Over Time. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management* (pp. 2335–2338). New York: ACM.



- McGregor, S.C. 2019. Social Media as Public Opinion: How Journalists Use Social Media to Represent Public Opinion. *Journalism*. <https://doi.org/10.1177/1464884919845458>.
- Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. 2019. *Reuters Institute Digital News Report 2019* (Vol. 2019). Reuters Institute for the Study of Journalism.
- Papacharissi, Zizi A. 2003. The Virtual Sphere: The Internet as a Public Sphere. *New Media and Society* 4 (1): 9–27.
- Papacharissi, Zizi A. 2010. *A Private Sphere: Democracy in a Digital Age*. New York: Wiley.
- Papacharissi, Zizi A., and Maria de Fatima Oliveira. 2012. Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication* 62 (2): 266–282. <https://doi.org/10.1111/j.1460-2466.2012.01630.x>.
- Stockmann, Daniela, and Teng Luo. 2017. Which Social Media Facilitate Online Public Opinion in China? *Problems of Post-Communism* 64 (3–4): 189–202.
- Trilling, D., P. Tolochko, and B. Burscher. 2017. From Newsworthiness to Shareworthiness: How to Predict News Sharing Based on Article Characteristics. *Journalism & Mass Communication Quarterly* 94 (1): 38–60.
- Vaccari, Cristian, Augusto Valeriani, Pablo Bonneau Barbera, Jost Rich, T. John, Jonathan Nagler, and Joshua Tucker. 2015. Political Expression and Action on Social Media: Exploring the Relationship between Lower- and Higher- Threshold Political Activities among Twitter Users in Italy. *Journal of Computer-Mediated Communication* 20: 221–239.
- Van der Brug, W. 2004. Issue Ownership and Party Choice. *Electoral Studies* 23 (2): 209–233.
- Van der Veer, Neil, Boeke, Steven, Hoekstra, Hans and Peters, Oskar. 2018. *Nationale Social Media Onderzoek 2018: Het Grootste Trendonderzoek van Nederland naar het Gebruik en Verwachtingen van Social Media #NSMO*. Newcom Research and Consultancy B.V.
- Wells, C., D.V. Shah, J.C. Pevehouse, J. Yang, A. Pelled, F. Boehm, J. Lukito, S. Ghosh, and J.L. Schmidt. 2016. How Trump Drove Coverage to the Nomination: Hybrid Media Campaigning. *Political Communication* 33: 669–676. <https://doi.org/10.1080/10584609.2016.1224416>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Marijn van Klingeren** is an Assistant Professor in Sociology at the Radboud University in Nijmegen. Her research interests include media patterns and media influences on public opinion, interethnic relations, migrant perceptions, and (Muslim) migrant integration processes.

**Damian Trilling** is an Associate Professor at the Department of Communication Science at the University of Amsterdam and affiliated with the Amsterdam School of Communication Research. Most of his research is related to the question how the changing media landscape affects the way in which citizens keep up with news and political information.

**Judith Möller** is an Assistant Professor at the Amsterdam School of Communication Research. In her research, she focuses on the effects of political communication, in particular social media and digital media.

