---

**Article:**
Galesic, M, Bruine de Bruin, W orcid.org/0000-0002-1601-789X, Dumas, M et al. (3 more authors) (2018) Asking about social circles improves election predictions. Nature Human Behaviour, 2 (3). pp. 187-193. ISSN 2397-3374

https://doi.org/10.1038/s41562-018-0302-y

---

# Asking about social circles improves election predictions

## Paper accepted for publication in Nature: Human Behaviour

M. Galesic[1,2,*], W. Bruine de Bruin[3-5], M. Dumas[1,6], A. Kapteyn[5], J.E. Darling[5,7], E. Meijer[5]

[1]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501.

[2]Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

[3]Centre for Decision Research, Leeds University Business School, Maurice Keyworth Building, Leeds LS2 9JT, UK.

[4]Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213.

[5]Center for Economic and Social Research, University of Southern California, 635 Downey Way, Los Angeles, CA 90089-3332.

[6]London School of Economics, Houghton St, London WC2A 2AE, UK.

[7]VHA Health Services Research and Development Center for the Study of Healthcare Innovation, Implementation, and Policy, VHA Greater Los Angeles Health Care System, Los Angeles, CA, USA

Election outcomes can be difficult to predict. A recent example is the 2016 U.S. presidential election, where Hillary Clinton lost five states that had been predicted to go for her, and with them the White House. Most election polls ask people about their own voting intentions: whether they will vote, and if so, for which candidate. We show that, compared to own-intention questions, social-circle questions that ask participants about the voting intentions of their social contacts improved predictions of voting in the 2016 U.S. and 2017 French presidential elections. Responses to social-circle questions predicted election outcomes on national, state, and individual levels, helped explain last-minute changes in people's voting intentions, and provided information about the dynamics of echo chambers among supporters of different candidates.

Past polls have asked people to indicate who they think will win the election, or to judge the probability that each candidate will win. Possibly because people know how their social contacts will vote,[1] such election-winner questions have successfully predicted many election outcomes[2,3]. However, election-winner questions have some imperfections. They do not straightforwardly predict actual vote shares because they ask for expectations that a candidate will win and not for the estimated percentage of voters who will vote for the candidate. They produce predictions on the national but not on the state and individual levels. Furthermore, they rely on people's inferences about the general population, which are likely influenced by sometimes inaccurate predictions reported in the media[4].

Social-circle questions can provide useful information in election polls for several reasons. It has been shown that people can provide relatively accurate judgments about various characteristics of their immediate social circles[5,6]. Averaged across a national sample, respondents' judged percentage of their social contacts with specific characteristics (such as having health problems) tend to come closer to the actual percentage in the general population than respondents' judged percentage of the population with these characteristics[7]. Moreover, reporting about friends' preferences for an unpopular candidate can be less embarrassing than admitting to personally having these preferences[8,9]. People's reports about their social contacts may also illuminate the social interactions that shape their beliefs and behaviours, and anticipate changes in own intentions over time due to social influence processes[10,11]. Finally, social-circle questions provide information about individuals who were not included in the sample of a

particular poll, thus implicitly increasing sample size and possibly reducing some of its sampling, nonresponse, and coverage bias[12].

We studied the usefulness of social-circle questions in two different elections: the 2016 U.S. presidential election, and the 2017 French presidential election. Held on November 9, the U.S. election essentially focused on two candidates - Hillary Clinton and Donald Trump. Other candidates were collectively not expected to win more than about 10% of the vote. In contrast, French elections involved at least five prominent candidates: François Fillon, Benoît Hamon, Marine Le Pen, Emmanuel Macron, and Jean-Luc Mélenchon, while also including six others with less prominence. In France, the election was held in two rounds: the first round occurred on April 23, with the second round on May 7 focusing on the top two candidates – Marine Le Pen and Emmanuel Macron.

In each country, we asked questions about participants' social contacts in two parts: (a) "What percentage of your social contacts are likely to vote in the upcoming election?" and (b) ("Of all your social contacts who are likely to vote, what percentage do you think will vote for [candidate]?"; (see Methods). In the U.S., we asked social-circle questions in two national surveys: the GfK election poll conducted in the week before the election[13] and the USC Dornsife/LA Times election poll conducted daily from July 2016 until after the election[14,15]. We compared the social-circle questions with two versions of own-intention questions: a standard version in the GfK poll, "If you were to vote in the presidential election that's being held on November 8th, which candidate would you choose?"[13], and a probabilistic version in the USC poll, "If you do vote in the election, what is the percent chance that you will vote for Clinton, Trump, or someone else?"[16,17]. Both were preceded by a question about going to vote at all (see Methods). In addition, the USC poll elicited election-winner expectations: "What is the percent chance that Clinton, Trump, or someone else will win?" In France, we asked the social-circle questions in the election poll conducted by survey research company BVA on a national sample in the week before the first round of the election. We compared the answers to social-circle questions with answers to own-intention questions of the form "Which candidate are you most likely to vote for?" (see Methods). These data allowed us to ask five research questions.

Our first research question examined whether asking about social circles improved predictions of national election results. Tables 1 and 2 summarize results from the U.S. and

French elections, including established measures of prediction error. In the U.S., social-circle questions were more accurate than own-intention questions, in predicting the whole distribution of vote shares for different candidates (Table 1 and Supplementary Information). We found lower values of error measures Mosteller 3[18] and $\bar{A}$[19,20] for social-circle than for own-intention questions. Compared to own-intention questions, social-circle questions predicted the difference between Clinton and Trump less well, as indicated by error measure Mosteller 5[18], which considers only the two main candidates. Which of these well-established error measures is more important will depend on the number of prominent candidates in a given election, and on the aims of the particular poll. In France, social-circle questions performed better than own-intention questions on all error measures and in both election rounds (Table 2). In both countries, social circle questions produced more accurate predictions of participation rates, with a particularly large improvement over own-intention questions in France. Possibly reflecting media forecasts of a substantial Clinton win, the USC poll's election-winner question erroneously predicted that Clinton would win, giving her a 53.4% chance compared to 42.5% for Trump and 4.1% for other candidates.

Second, we investigated whether social-circle questions improved predictions of state election results. Social-circle questions produced more accurate predictions of state winners, as compared to own-intention questions, as seen in both US polls. Consequently, social-circle questions predicted the number of electoral votes for each candidate better than own-intention questions (Table 1 and Supplementary Information). USC's social-circle questions were the only ones that predicted Trump winning the majority of electoral votes. Moreover, GfK's and USC's polls achieved above-chance accuracy in predicting winners at the state level, especially with social circle questions. These results were obtained despite sample sizes of only 27 participants per state for GfK's polls, and 44 per state for USC's polls. GfK's and USC's polls respectively predicted 67% and 77% of states correctly with social-circle questions, as compared to 65% and 61% with own-intention questions. For further comparison, aggregates of 3,073 state polls (including 60 polls per state on average) predicted 90% of states correctly[21]. Social-circle questions were particularly useful for predicting election outcomes in a priori defined "swing states"[22] (CO, FL, IA, MI, NC, NV, NH, OH, PA, VA, and WI). For GfK and USC, social circle questions respectively predicted 82% and 73% of swing states correctly, while own-intention questions accurately predicted 46% and 64% of swing states correctly, and aggregates of 3,073

state polls accurately predicting 55% of swing states correctly. Social-circle questions were also more successful than both own-intention questions and aggregate polls in predicting winners of the five swing states that unexpectedly went to Trump (FL, MI, NC, PA, and WI). They predicted four of these states correctly, compared to three by own-intention questions and zero by aggregate polls. In sum, these results suggest that people possess valuable information about their social circles, which could be used to improve election predictions at national and state levels.

< TABLES 1 AND 2 ABOUT HERE >

Third, we examined whether social-circle questions benefited predictions of individual voting behaviour. We found that, over time, changes in social-circle reports predicted subsequent changes in own voting intentions. For participants who completed USC surveys in August, September, late October/early November, and immediately after the election (N=1,263), social-circle questions contributed to the explanation of their actual voting behaviour over and above own-intention questions. Fig. 1A shows that, up until the week before the election, participants reported that they were on average more likely to vote for Clinton than for Trump, while more ended up voting for Trump than for Clinton. Fig. 1B shows a reversal toward Trump in social-circle reports as early as September 2016, when own-intention questions were still predicting a lead for Clinton. A weighted average of own intentions and social-circle estimates led to more accurate predictions of individual voting behaviour than own intentions alone (with weights being regression coefficients in a model including both types of questions, see Extended Data Table 1). Of note, election-winner expectations did not contribute to explanations of voting behaviour over and above own intentions and social-circle questions (Extended Data Table 1). Similar patterns were observed throughout the pre-election period, with own intentions and social-circle reports jointly contributing to explanations of own intentions in subsequent survey waves (Extended Data Tables 2a and 2b).

< FIGURE 1 ABOUT HERE >

Fourth, we analysed whether social-circle questions helped to explain last-minute changes in voting intentions. Not all participants ended up voting for the candidate they announced as their favourite in the week before the election (Extended Data Table 3). For example, participants whose own intentions mismatched those in their social circles were less

likely to eventually vote for their intended candidate (Extended Data Fig. 1). While some of these participants had less strong intentions to vote for their preferred candidate in the first place, our overall results suggest that social-circle reports foretold a switch in voting intentions before it happened. Generally, changes in participants' social circles over time predicted their later intentions to vote for specific candidates and to vote at all, as revealed by vector autoregression modelling and Granger causality tests[23,24] (Fig. 1C, Extended Data Tables 4 and 5). This pattern of results was found for both Trump and Clinton voters, suggesting that participants' perceptions of how social contacts would vote affected their own beliefs regarding the candidates. Additionally, Trump voters appeared to influence later changes in their social circles while Clinton voters did not.

Our final research question was whether asking about social-circles provided insights about the dynamics of echo chambers. Social-circle questions revealed increased homogenization of Trump voters' social circles over time. Fig. 2 shows the percentage of like-minded social contacts that Trump and Clinton voters reported in the USC poll. Extreme echo chambers would be seen in social circles that include nearly 100% like-minded individuals. In August 2016, individuals who eventually voted for Trump and those who eventually voted for Clinton had similarly diverse social circles. Respectively, their social circles included on average around 68% and 71% like-minded individuals. However, over time, social circles of Trump voters included increasingly more like-minded individuals. In contrast, we did not observe a similar increase in the homogeneity of Clinton voters' social circles. Hence, the additional Trump voters were likely coming from people who previously did not plan on voting, were undecided, or were planning to vote for third candidates. It is also possible that Trump voters were more inclined to exclude Clinton supporters from their social circles than were Clinton voters to exclude Trump fans. In any case, the homogenization continued after the election, when Trump voters reported social circles consisting of on average 77% like-minded individuals, compared to 68% among Clinton voters. Just after the election, 42% of Trump voters had social circles that included 90% or more like-minded individuals, compared to only 30% of such participants among Clinton voters. When further investigating whether the homogeneity of social circles was related to sociodemographic variables (Extended Data Table 6 and Extended Data Figs. 2 and 3), we found moderate relationships with participants' political leanings, age, education, and U.S. state of residence. For Trump voters, homogenization of social circles was

particularly pronounced among older voters aged 65 or older, in particular in states that voted Republican. Education played an additional role, with less educated Trump voters homogenizing more and faster than more educated ones. In comparison, age did not predict homogenization for Clinton voters. In addition, in strongly Democrat states, more educated Clinton voters had somewhat more homogeneous circles than less educated ones (see Extended Figures 2 and 3).

< FIGURE 2 ABOUT HERE >

Taken together, our results make two contributions. First, people's reports about their social circles can improve predictions of election results and enhance understanding of individual voting behaviour. We observe these findings across different poll designs in two countries with different political systems, suggesting that other election polls could potentially benefit from including social-circle questions. Social-circle questions may also be useful in surveys aiming to forecast other beliefs and behaviours. One reason for the usefulness of social-circle questions could be the increased implicit sample size, which was reflected in reduced standard errors of social-circle compared to own-intention questions. For GfK's poll, standard errors for predictions from social circles vs. own intentions were 0.78 vs. 1.22 for Clinton, 0.78 vs. 1.21 for Trump, and 0.31 vs. 0.77 for others. Similarly, USC poll standard errors for predictions from own intentions were 0.92 vs. 1.35, 0.94 vs. 1.45, and 0.35 vs. 0.75. Social-circle reports might provide information about people who would otherwise be missing from polls due to coverage, sampling, or nonresponse errors[12]. Social-circle questions could therefore be particularly useful when polls must rely on relatively small samples in some states. Another reason for the usefulness of social-circle might be that participants who are reluctant to report that they favour a potentially embarrassing option could nevertheless be willing to report that their social circle favours it[8,9]. Finally, through processes of social influence, individuals' voting intentions could indeed become more similar to the prevailing opinion in their social circles over time[10,11].

Our second main finding is that asking about social circles can provide insights into the social dynamics that shape individual voting behaviour. We find interesting differences between Trump voters and Clinton voters. Trump voters seemed to be influenced by their peers and influencing them in turn (Fig. 1C and Extended Data Table 5). Clinton voters appeared to be mostly influenced by others while not influencing others themselves. One possible explanation for this finding is that Trump voters might have been more likely to project their own intentions

onto intentions of their peers, perceiving them as more similar to themselves than they were[25]. It is also possible that they were influencing their friends and family to vote for Trump, or that the composition of their social circles was changing over time to include more Trump supporters. These differences between Trump and Clinton voters are echoed by the finding of increased homogenisation of social circles of Trump, but not Clinton voters (Fig. 2). This pattern of homogenisation likely results from several inter-related processes. One is Trump supporters' increasing suspicion of the "mainstream media"[26] and greater reliance on in-group information sources. Another is "unfriending" of people with incompatible political opinions, practiced by supporters of both candidates[27]. Perceived homogeneity can further increase if people are reluctant to disclose political views that are not in accord with the prevailing opinion among their peers[28]. Our results are in line with a recent analysis of Twitter data that showed significant homogeneity and isolation of Trump voters relative to supporters of other candidates[29].

Overall, social-circle questions are a way of tapping into the "local" wisdom of crowds[30-32]. Standard election-winner questions attempted to tap into the wisdom of crowds by asking people about their predictions for overall election results[2-4]. This can be problematic because people do not have a direct experience with everyone in the general population. Instead, they have to make population inferences based at least in part on second-hand information, such as sometimes erroneous predictions reported in the media. In contrast, social-circle questions harvest people's direct experiences with their immediate social environments[5-7,33]. It is important to note that survey sampling design will affect the usefulness of social-circle questions. If social-circle reports come from a biased, non-representative sample of the overall population, their average will likely be a biased estimate of true population values. In well-designed samples of the population of interest social-circle questions can improve survey estimates, especially when these are otherwise based on small samples or when they pertain to socially sensitive beliefs and behaviours. In addition, social-circle reports can provide valuable information about social interactions that shape individual beliefs and behaviours.

**Methods:**

**1. Aggregate polls**

For election predictions based on aggregate polls in the U.S., we used data from 1,106 national polls[21] and 3,073 state polls, summarized by the site FiveThirtyEight.com from [34]. In

France, we used results of 20 different polls conducted in the week before election round 1, and 18 before round 2[35].

## 2. Individual polls

We investigated the usefulness of social-circle questions in three individual polls, described below. All participants gave informed consent. The research was approved by the University of Southern California's Dornsife's Institutional Review Board (USC poll), and the Federalwide Assurance Signatory Official of the Santa Fe Institute (all polls).

### 2.1. USC Dornsife/LA Times Presidential Election Poll

#### Question texts

Introduction: In this interview, we will ask you questions about the upcoming general election for President of the United States on Tuesday November 8, 2016. All questions ask you to think about the percent chance that something will happen in the future. The percent chance can be thought of as the number of chances out of 100. You can use any number between 0 and *100. For example, numbers like: 2 and 5 percent may be 'almost no chance', 20 percent or so may mean 'not much chance', a 45 or 55 percent chance may be a 'pretty even chance', 80 percent or so may mean a 'very good chance', and a 95 or 98 percent chance may be 'almost certain'.*

Own-intention questions: (a) What is the percent chance that you will vote in the Presidential election? (b) If you do vote in the election, what is the percent chance that you will vote for Clinton? And for Trump? And for someone else? Order of candidates was randomized for this and other questions in all three polls.

Social-circle questions: Now we would like you to think of your friends, family, colleagues, and other acquaintances of 18 years of age or older that you have communicated with at least briefly within the last month, either face-to-face, or otherwise. We will call these people your social contacts. (a) What percentage of your social contacts are likely to vote in the upcoming election for President? For instance, 0% means that you think none of your social contacts will vote, and 100% means that all of your social contacts will vote. If you are not sure, just try to give your best guess. (b) For the next question, please consider only those of your social contacts who are likely to vote in the upcoming election for U.S. President. Of all your

social contacts who are likely to vote, what percentage do you think will vote for Clinton, Trump, or someone else? For instance, 0% would mean that you think no voters in your social circle will vote for that candidate, and 100% means that all voters in your social circle will vote for that candidate. Again, if you are not sure, just try to give your best guess.

Election-winner expectations questions: What is the percent chance that Clinton will win? And Trump? And other candidates?

**Sample**

Participants were members of the Understanding America Study at the University of Southern California's Dornsife Center for Economic and Social Research. This longitudinal study[36] included close to 6,000 U.S. residents who were randomly selected from among all households in the United States using address-based sampling. They were recruited by a combination of mail, phone, and web surveys. Members of recruited households who did not have Internet access were provided with tablets and Internet service. In May 2016, all panel members who were U.S. citizens were asked to respond to a pre-election survey. Those who completed the study and agreed to participate constituted the election poll panel.

Starting from July 4, 2016, each member of the poll panel was invited to answer the election poll once a week[37]. Members received the invitation to participate each week on the same day of the week, but they were allowed to respond up to 6 days later (i.e., until the day before the next invitation). On average across waves, study completion rates were 70%. As reported by UAS[37], the average panel recruitment rate, reflecting those individuals who completed the initial mail survey among those who consented to participate in the UAS, was 29.7%. The percentage of active panel members was 13.6%[36]. Combined with the study completion rate, the cumulative response rate for the studies reported here was 9.5%.

Five study waves asked all three types of questions of interest for this study (own-intention, social-circle, and election-winner questions): (1) July 11–23 (N=1,782), (2) August 8–20 (N=2,726), (3) September 12–24 (N=2,882), (4) October 31–November 7 (N=2,240), and (5) after the election, November 9–21 (N=3,798). In all waves except Wave 4, all questions were asked together, In Wave 4 only, social-circle questions were asked in a separate questionnaire from own-likelihood and election-winner questions. That is, social-circle questions were asked starting from November 3, and all participants completed all questions within a window of about

3 days. The pattern of results presented in the main text does not change if we analyse only the 969 participants who, in Wave 4, completed all questions on the same day. This would be expected because all pre-election "surprises" occurred before this survey period, with the last being the October 28 FBI announcement that they were re-opening the investigation of Clinton's emails.

In this paper, we analysed two subsamples of participants: those who completed Wave 4 (excluding a small number of participants who did not answer all questions, resulting in the total N=2,229), and those who completed each of the Waves 2, 3, 4, and 5 (N=1,263).

**Analyses**

Survey weights were constructed by a raking procedure that matched the sample to national population benchmarks based on the May 2016 Current Population Survey age by sex, race/ethnicity, sex by education, and household size by income. An additional weighting variable, reflecting whether or not participants voted in the 2012 election, was used to achieve representative proportions of voters for different candidates[38]. The weights were used only for the analyses on N=2,229 individuals who participated in the last wave before the election (results shown in Table 1). The analyses on N=1,263 individuals who participated in all survey waves from August to after the election were done on unweighted data, because the goal of these analyses was to describe that particular sample and not to make inferences about the overall population (Results in Figs. 1, 2, and Extended Data Figures and Tables).

In line with previous studies using probabilistic questions about voting behaviour[16,17], USC predictions of election outcomes were derived by (1) multiplying each participant's own (or social-circle's) likelihood to vote by his/her (or his/her social-circle's) likelihood to vote for each of the candidates and (2) estimating the ratio of the resulting variable and the average of the participant's own (or social-circle's) likelihood to vote across all participants[37].

**2.2. GfK Election Survey**

**Question texts**

Own-intention questions: (a) How likely are you to vote in this upcoming election? (b) If *you were to vote in the presidential election that's being held on November 8th, which candidate* would you choose? Response options were Clinton, Trump, Johnson, Stein, another candidate,

undecided, and would not vote. The question was slightly modified for those who were certain to vote/had already voted: *Thinking about the presidential election that's being held on November* 8th, for whom will/did you vote? Likelihood to vote was additionally determined by asking whether participants were registered voters in their state of residence, and whether they had voted in previous elections.

Social-circle questions: Same as for the USC poll.

**Sample**

Participants were selected from GfK's national, probability-based online KnowledgePanel, which currently includes 55,000 active members[39]. They were primarily recruited using address-based sampling methods, including telephone follow-up for refusal conversions. Adults who were selected to join KnowledgePanel but did not have access to the Internet were provided with Internet access and a web-based device at no cost. For this study, the KnowledgePanel sample included active panel members who were 18 years or older and lived in the United States at the time of the study. Participants were selected using a proprietary probability proportional to size (PPS) sample algorithm. As a result, the final sample reflected the demographic profile of adults 18 years or older based on targets derived from the March 2016 Current Population Survey. The sample was also balanced in respect to party identification (Democrats, Republicans, and Independent/Others) as measured on an earlier panel profile survey, with target proportions based on the average values obtained from eight different probability-based national polls fielded in the two months prior to this study.

A total of 4,181 members of KnowledgePanel were included here. The field period was November 4, 2016 (1:30 a.m. EST) to November 8, 2016 (11:45 a.m. EST). Of those who were invited, 2,367 members completed the survey (a 56.5% study completion rate). As reported by GfK[40], the average panel recruitment rate for participants in this study was 13.0%. Of the recruited households, 62.4% completed the initial profile survey. Together with study completion rate, this leads to a cumulative response rate of 4.6%.

**Analyses**

Standard geodemographic weights were computed for all participants, regardless of voter registration and likelihood to vote, using iterative proportional fitting or raking. National

population benchmarks based on March 2016 Current Population Survey data were used to create weighting targets based on region, age by sex, education, income, and race/ethnicity[40].

Predictions based on own likelihood to vote for different candidates, shown in Table 1 and Fig. 1, were weighted answers to this question for the subset of participants who were likely voters. These were determined by self-reports as all who indicated being registered voters in the state of their residence, were definitely likely to vote or had already voted, or said they would probably vote and also indicated they always or almost always voted in elections. In all, 1,897 of the respondents were determined to be likely voters (80.1% of all participants). Of those, 1,822 answered both the question about own likelihood to vote for different candidates and the questions about social-circle likelihood to vote. Predictions based on social-circle likelihood to vote for different candidates were obtained as described in the section about USC methodology above.

### 2.3. BVA French Presidential Election Poll

#### Question texts

Own intentions: Participants were asked about their voting intentions in the first and the second round of the election using the standard BVA methodology: *During the first round of the presidential election, which candidate are you most likely to vote for? (Lors du premier tour de l'élection présidentielle, quel serait le candidat pour lequel il y aurait le plus de chance que vous votiez?)* Response options were the 11 candidates, as well as *I will not go vote* (used to infer participation rates), and *I will vote blank*. For the second round, participants were asked: *Here is the list of candidates who, according to polls, should include the 2 qualified for the second round. Could you indicate how you would rank each of them?* (*Voici la liste des candidats parmi lesquels, d'après les sondages, devraient se trouver les 2 qualifiés du second tour. Pourriez-vous indiquer, dans l'ordre de vos préférences, les candidats…*) Response options were the four top candidates and *None of them*.

Social-circle questions for the first round of election asked: (a) According to you, what share of your social circle will go vote in the first round of the election? (A votre avis, quelle est *la part de votre entourage qui ira voter au premier tour de l'élection?)* and (b) Amongst the members of your social circle who should go vote in the first round of the presidential election, how do you expect their votes to be distributed between the different candidates? (Parmi les

*membres de votre entourage qui devrait aller voter au premier tour de l'élection présidentielle,* comment devraient se répartir les votes en faveur des différents candidat?) The options were Dupont-Aignan, Fillon, Hamon, Le Pen, Macron, Mélenchon, other candidates, and voting blank. For the second round, participants were asked: (a) "Suppose that Emmanuel Macron and Marine Le Pen are the candidates in the second round. What will be the share of your social *circle that will go vote in the second round?" (Supposons qu'Emmanuel Macron et Marine Le* Pen soient les candidats du second tour. Quelle est la part de votre entourage qui ira voter au second tour?) and (b) a similar question as in the first round focused on only Le Pen, Macron, and not voting for either of them.

**Sample**

In line with standard French polling practices[41], the sample was selected from the BVA online access panel by quota sampling. The quotas were designed to represent the French population by gender, age, partisan affiliation, employment, region, and settlement size, following the guidelines of the French National Statistical Institute. Only registered voters were contacted. The survey took place from April 17 to 22, 2017, just before the first round of election on April 23. According to BVA, of 1,685 people who satisfied the quota and were invited to participate, 59.5% completed the study, for the final sample of 1,003 participants.

**Analyses**

Post-stratification weights were used to adjust the sample frequencies to the general population according to gender, age, employment, region, and settlement size.

Predictions based on own likelihood to vote, shown in Table 2, were weighted answers of all participants (who were all registered voters, by design) to the questions about intention to vote for different candidates, to vote blank, or to not vote, in the first and second round. Predictions based on social circle questions were obtained as described in the section about USC methodology above, using answers to questions about the percentage of social circle who will not vote or will vote blank, and who will vote for different candidates among those social contacts who will vote.

**Data availability**

The data that support the findings of this research are available from the corresponding author upon request.

## Code availability

Stata and SPSS codes for all analyses are available from the corresponding author upon request.

## References


1. Rothschild, D. M. &Wolfers, J., Forecasting elections: Voter intentions versus expectations. https://ssrn.com/abstract=1884644. (2011).
2. Lewis-Beck, M. S. & Skalaban, A. Citizen forecasting: Can voters see into the future? Brit. J. Polit. Sci. **19**, 146-153 (1989).
3. Graefe, A. Accuracy of vote expectation surveys in forecasting elections. Public Opin. Q. **78**, 204-32 (2014).
4. Irwin G. A. Van Holsteyn J. J. M. According to the polls: The influence of opinion polls on expectations. Public Opin. Q. 66, 92-104 (2002).
5. Nisbett, R. E. & Kunda, Z. Perception of social distributions. J. Pers. Soc. Psychol., **48**, 297-311 (1985).
6. Dawtry, R. J., Sutton, R. M. & Sibley, C. G. Why wealthier people think people are wealthier, and why it matters: From social sampling to attitudes to redistribution. Psychol. Sci. **26**, 1389-1400 (2015).
7. Galesic, M., Olsson, H. & Rieskamp, J. Social sampling explains apparent biases in judgments of social environments. Psychol. Sci. **23**, 1515-1523 (2012).
8. Sudman, S. & Bradburn, N. M. Asking Questions: A Practical Guide to Questionnaire Design (Jossey-Bass, 1982).
9. J. Darling, A. Kapteyn, Hidden Trump Voters? (AAPOR, 2017).
10. Huckfeldt, R. R. & Sprague, J. Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign (Cambridge U. Press, 1995).
11. Sinclair, B. The Social Citizen: Peer Networks and Political Behavior (U. of Chicago Press, 2012).
12. Groves, R. Survey Errors and Survey Costs (Wiley, 2004).
13. GfK. 2016 Election Survey. http://www.gfk.com/products-a-z/us/public-communications-and-social-science/2016-election-survey (2016).

14. USC Dornsife/LA Times. Presidential Election "Daybreak" Poll—Understanding America Study. http://cesrusc.org/election/ (2016).
15. Kapteyn, A., Meijer, E. & Weerman, B. Methodology of the RAND Continuous 2012 Presidential Election Poll. Working Paper No. WR-961 (RAND Corporation, 2012).
16. Delavande, A. & Manski, C. F. Probabilistic polling and voting in the 2008 presidential election: Evidence from the American Life Panel. Public Opin. Q. **74**, 433-459 (2010).
17. Gutsche, T. L., Kapteyn, A., Meijer, E., & Weerman, B. The RAND Continuous 2012 Presidential Election Poll. Public Opin. Q. **78**, 233-254 (2014).
18. Mosteller, F. & Doob, L. W. The Pre-Election Polls of 1948. (Social Science Research Council, 1949).
19. Martin, E. A., Traugott, M. W. & Kennedy, C. A review and proposal for a new measure of poll accuracy. Public Opin. Q. **69**, 342-369 (2005).
20. Arzheimer, K. & Evans, J. A new multinomial accuracy measure for polling bias. Polit. Anal. **22**, 31-44 (2014).
21. FiveThirtyEight.com. National Polls. https://projects.fivethirtyeight.com/2016-election-forecast/national-polls/ (2016).
22. FiveThirtyEight.com. Election Update. https://fivethirtyeight.com/features/election-update-the-state-of-the-states/ (2016).
23. Abrigo, M. R. & Love, I. Estimation of panel vector autoregression in stata: a package of programs. Stata J. **16**, 778-804 (2016).
24. Holtz-Eakin, D., Newey, W. & Rosen, H. S. Estimating vector autoregressions with panel data. Econometrica **56,** 1371-1395 (1988).
25. Dawes, R. Statistical criteria for establishing a truly false consensus effect. J. Exp. Soc. Psychol. **25**, 1-17 (1989).
26. Gallup.com. Americans' Trust in Mass Media Sinks to New Low. http://www.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx (2016).
27. Duggan, M. Smith, A. & Pew Research Center. The tone of social media discussions around politics. http://www.pewinternet.org/2016/10/25/the-tone-of-social-media-discussions-around-politics/ (2016).
28. Noelle-Neumann, E. J. Commun. **24**, 43-51 (1974).

29. Thompson, A. Parallel Narratives. https://news.vice.com/story/journalists-and-trump-voters-live-in-separate-online-bubbles-mit-analysis-shows (2016).
30. Tetlock, P. E. & Gardner, D. Superforecasting: The art and science of prediction (Random House, 2016).
31. Prelec D. A Bayesian truth serum for subjective data. Science **306**, 462-6 (2004).
32. Prelec D, Seung H.S., McCoy J. A solution to the single-question crowd wisdom problem. Nature, **541**, 532-5 (2017).
33. Trafalgar Group. Press release 11/09. http://us13.campaign-archive1.com/?e=9941a03e2e&u=99839c1f5b2cbb6320408fcb8&id=f7b7326d0b (2016).
34. FiveThirtyEight.com. Alabama—2016 Election Forecast. https://projects.fivethirtyeight.com/2016-election-forecast/alabama/ (2016).
35. Opinion polling for the French presidential election, 2017. https://en.wikipedia.org/wiki/Opinion_polling_for_the_French_presidential_election,_2017 (2017).
36. USC Dornsife. Sample and Recruitment - Understanding America Study. https://uasdata.usc.edu/index.php (2016).
37. Meijer, E. Sample Selection in the Daybreak Poll. http://cesrusc.org/election/samp-estim01.pdf (2016).
38. Meijer, E. Weighting the Daybreak Poll. http://cesrusc.org/election/weights03.pdf (2016).
39. GfK. Knowledge Panel. http://www.gfk.com/products-a-z/us/knowledgepanel-united-states/ (2016).
40. Thomas, R. K., Barlas, F. M., Weber, A. & McPetrie, L. Election 2016 Survey—National Election Omnibus (GfK Custom Research, 2016).
41. Norpoth, H. A win for the pollsters: French election predicted accurately. http://thehill.com/blogs/pundits-blog/international-affairs/330272-after-rough-2016-pollsters-get-it-right-in-french (2017).
42. Cousineau, D. Tutorials Quant. Methods Psychol. **1**, 42-45 (2005).

**Supplementary Information:** Table with detailed results for the U.S. elections is available online (TableSI_fin.xls).

**Acknowledgments**: The project described in this paper relies in part on data from survey(s) administered by the Understanding America Study (UAS), which is maintained by the Center for Economic and Social Research (CESR) at the University of Southern California (USC). The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of USC or UAS. This project was supported in part by grants from the National Science Foundation (MMS-1560592), the Swedish Foundation for the Humanities and the Social Sciences (Riksbankens Jubileumsfond) Program on Science and Proven Experience. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. We thank Henrik Olsson for helpful comments on a preliminary version of the paper, and Anita Todd for editing the manuscript.

**Author Contributions:** M.G. and W.B.B. designed the research questions and the social circle measures. A.K., J.E.D., and E.M. designed data collection methods and collected data within the USC study. M.G. and E.M. analysed the U.S. data. M.D. translated and adjusted the materials for data collection in France and analysed the French data. All authors contributed to the writing of the paper.

**Competing Interests:** The authors declare no competing financial or non-financial interests.

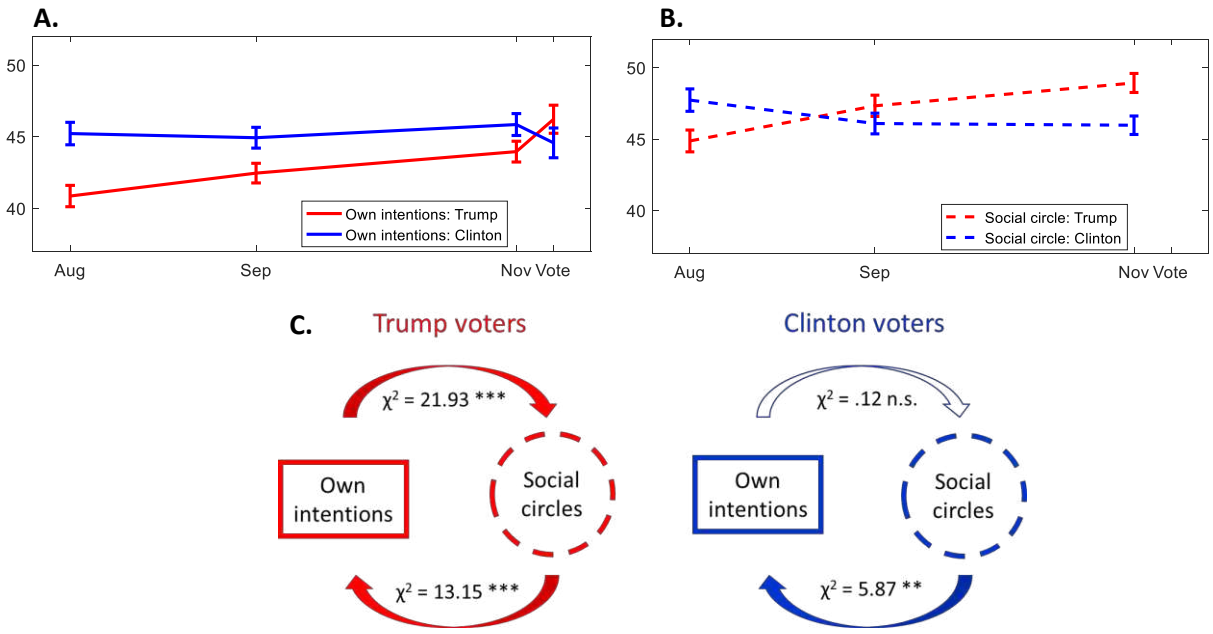



**Figure 1.** Shifts in average individual voting intentions and behaviour (A) were announced by shifts in social circles (B). Error bars in (B) show within-subjects 95% confidence intervals[42]. Granger causality tests in (C) suggest that social circles influenced own intentions reported weeks later. For Trump voters, own intentions also appeared to influence subsequent social-circle reports. Results are for N= 1,263 individuals who participated in USC's survey waves in August, September, early November, and immediately after the election. Because we are interested in predictions of individual behaviour in this particular sample, estimates are adjusted for likelihood of voting (see Methods) and otherwise unweighted.

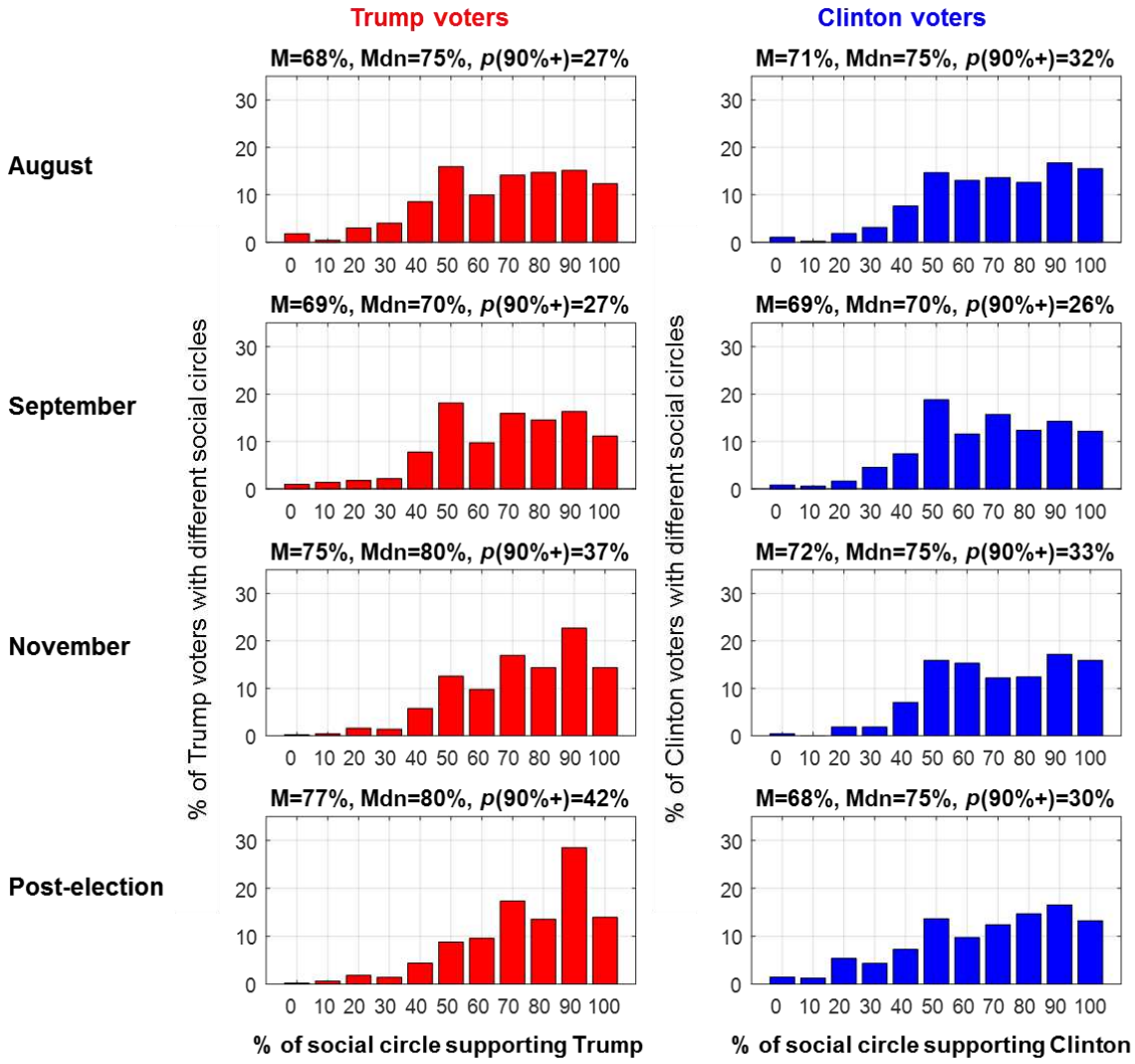



**Figure 2.** The extent of "echo chambers" among Clinton and Trump voters, over time. Data are for N= 1,263 individuals who participated in all four study waves. Shown are unweighted proportions of participants' social circles that they reported would vote (or had voted) for their preferred candidate.

**Table 1.** Actual results of the U.S. 2016 presidential election, predictions based on survey questions, and indicators of predictions' accuracy

| | Actual election outcome | Aggregate polls | GfK poll | | USC poll | |
|---|---|---|---|---|---|---|
| | | Own intention | Own intention | Social circle | Own intention | Social circle |
| Participation rate | 54.8 | - | 76.5 | 72.8 | 80.4 | 76.4 |
| % of popular vote | | | | | | |
| Clinton | 48.2 | 45.7 | 46.2 | 50.2 | 44.8 | 45.5 |
| Trump | 46.1 | 41.8 | 43.2 | 43.7 | 46.3 | 49.4 |
| Other | 5.7 | 12.5 | 10.6 | 6.1 | 8.9 | 5.1 |
| Electoral votes to Clinton (based on state-level predictions of popular vote) | 232 | 323 | 298 | 293 | 305 | 258 |
| **Error measures for national predictions (state-level predictions) of popular vote** | | | | | | |
| Error of predicted difference between 2 main candidates (Mosteller 5) | | 1.8 (1.3) | 0.9 (1.5) | 4.4 (5.0) | -3.6 (-0.6) | -6.0 (-1.4) |
| Average absolute error of predicted vote share for all candidates (Mosteller 3) | | 4.5 (2.1) | 3.3 (7.1) | 1.6 (3.9) | 2.3 (6.6) | 2.2 (5.9) |
| Average absolute log ratio of predicted and actual odds for all candidates ($\bar{A}$) | | 0.38 (0.14) | 0.29 (0.47) | 0.08 (0.24) | 0.21 (0.45) | 0.12 (0.35) |

Note: Results of the GfK poll were based on a probabilistic national sample of N= 1,822 participants interviewed from November 3 to the morning of November 8. Results of the USC poll were based on a probabilistic national sample of N= 2,229 participants interviewed from October 31 to November 7. For error measures, lower absolute values are better. For aggregate polls, question wording varied. In GfK's poll, own-intention questions asked which candidate participants would vote for, and in USC's poll they asked participants to judge the percent chance of voting for each candidate. Comparison with actual election results as well as with aggregate results of 1,106 national polls (for predictions of popular vote) and 3,073 state polls (for predictions of electoral votes) as summarized at fivethirtyeight.com[21,34] suggest that both GfK and USC polls have satisfactory accuracy. Note that Clinton eventually received 227 and Trump 304 electoral votes, because some electors have defected.

**Table 2.** Actual and predicted results of the French 2017 presidential election, and indicators of predictions' accuracy

| | Election round 1 | | | | Election round 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | Actual election results | Aggregate polls: Own intention | BVA poll: Own intention | BVA poll: Social circle | Actual election results | Aggregate polls: Own intention | BVA poll: Own intention | BVA poll: Social circle |
| Participation rate* | 75.8 | 73.2 | 89.6 | 74.7 | 66.0 | 74.0 | 81.3 | 68.7 |
| % of popular vote | | | | | | | | |
| Macron | 24.0 | 23.8 | 25.9 | 24.6 | 66.1 | 60.8 | 62.3 | 64.2 |
| Le Pen | 21.3 | 22.3 | 22.3 | 21.8 | 33.9 | 39.2 | 38.5 | 35.8 |
| Fillon | 20.0 | 19.5 | 15.2 | 17.3 | | | | |
| Mélenchon | 19.6 | 18.9 | 19.7 | 19.6 | | | | |
| Hamon | 6.4 | 7.6 | 7.3 | 8.8 | | | | |
| Others | 8.7 | 7.9 | 9.6 | 7.9 | | | | |
| **Error measures** | | | | | | | | |
| Error of predicted difference between main candidates (Macron & Le Pen, Mosteller 5) | | -1.2 | 0.8 | 0.1 | | -10.6 | -8.4 | -3.9 |
| Average absolute error of predicted vote share for all candidates (Mosteller 3) | | 0.8 | 1.6 | 1.2 | | 5.3 | 4.2 | 1.9 |
| Average absolute log ratio of predicted and actual odds for all candidates ($\bar{A}$) | | .08 | .12 | .11 | | .23 | .18 | .09 |

Note: Results of the BVA poll were based on a national quota sample of N=1,003 participants interviewed from April 17-22, 2017. For error measures, lower values are better. For comparison, we provide actual election results, as well as aggregate polls results based on questions about own voting intentions asked in 20 different polls in a week before election round 1, and 18 before round 2[35]. Note that compared to aggregate polls own-intention predictions from the BVA poll has satisfactory accuracy. Within the BVA poll, social-circle predictions always outperform those based on own-intention ones. *Non-participation count includes people who did not vote as well as those who casted blank ballots.

# Supplementary Information

# Asking about social circles improves election predictions

M. Galesic[1,2,*], W. Bruine de Bruin[3-5], M. Dumas[1,6], A. Kapteyn[5], J.E. Darling[5,7], E. Meijer[5]

[1]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501.

[2]Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

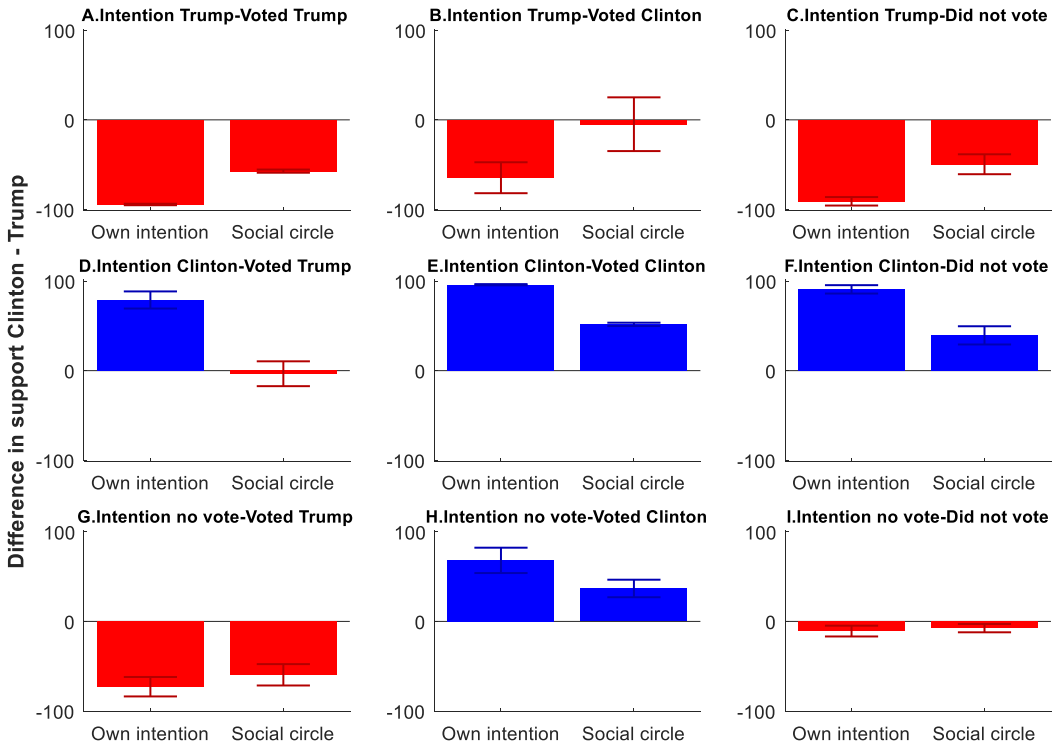
[3]Centre for Decision Research, Leeds University Business School, Maurice Keyworth Building, Leeds LS2 9JT, UK.

[4]Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213.

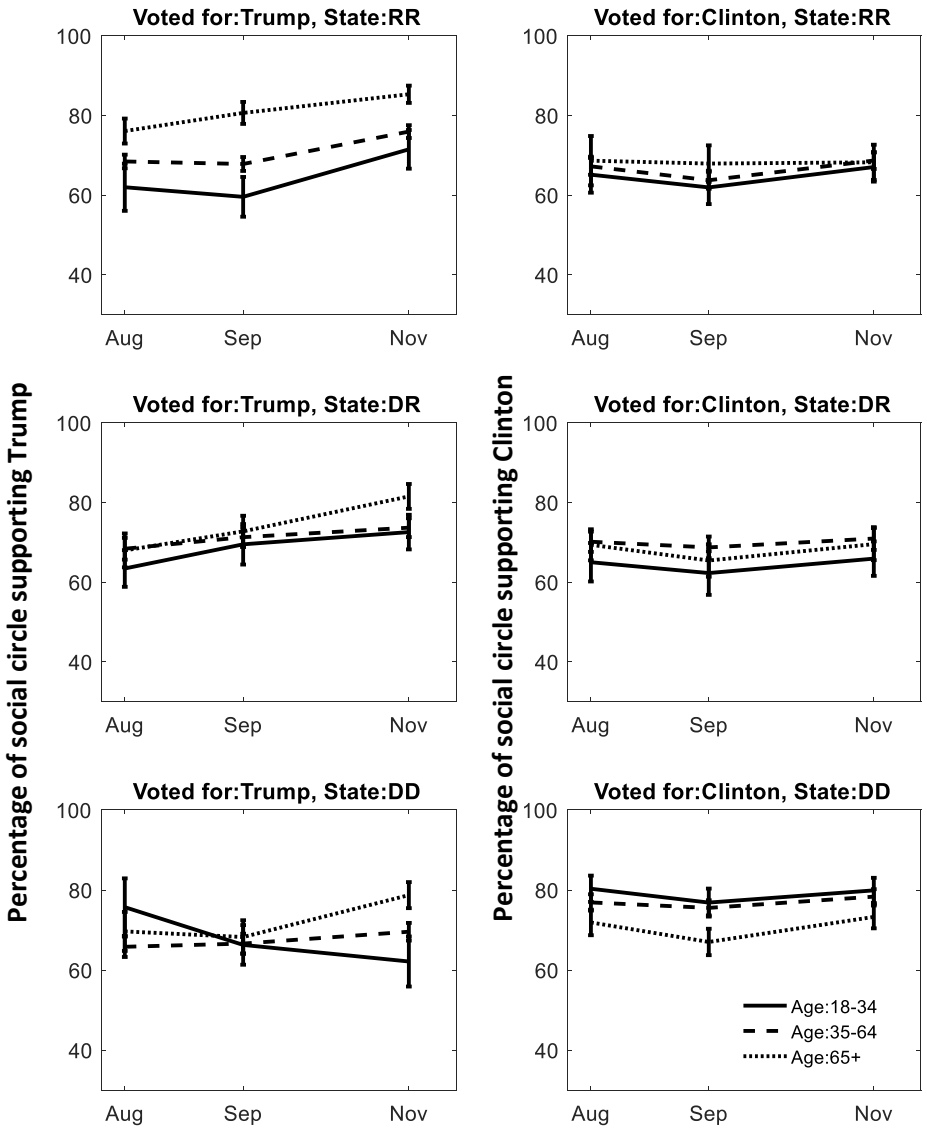
[5]Center for Economic and Social Research, University of Southern California, 635 Downey Way, Los Angeles, CA 90089-3332.

[6]London School of Economics, Houghton St, London WC2A 2AE, UK.

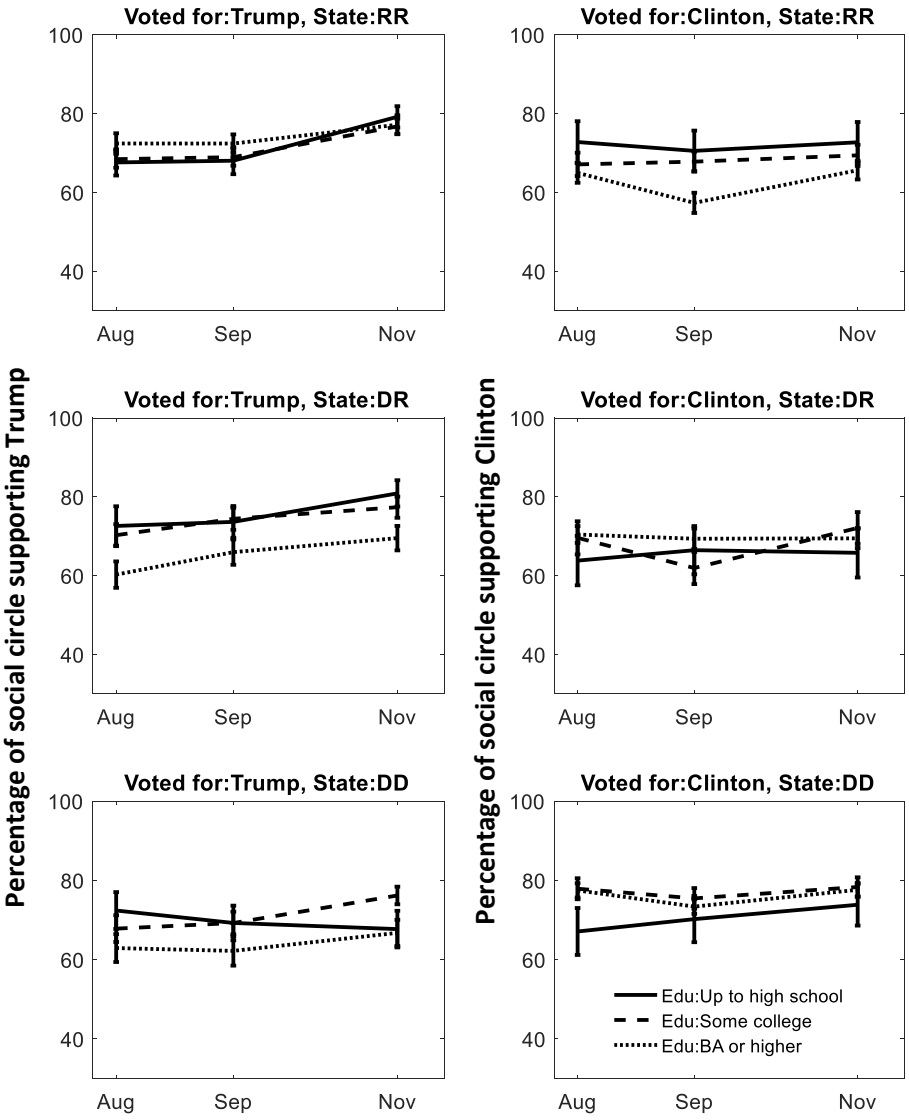
[7]VHA Health Services Research and Development Center for the Study of Healthcare Innovation, Implementation, and Policy, VHA Greater Los Angeles Health Care System, Los Angeles, CA, USA




**Extended Data Figure 1.** Own intentions and social-circle reports for nine different groups of participants, defined by the correspondence of their voting intentions a week before the election and their actual voting behaviour. Participants whose actual voting behaviour was in line with their stated intentions had social circles with higher percentage of individuals intending to vote for the same candidate (panels on the diagonal, A, E, and I). In contrast, participants who switched from their intended candidate to another candidate (panels B and D) had more heterogeneous social circles, equally likely to support both main candidates. Also interesting are participants who said they would not vote but who did vote (panels G and H). Both they and their social circles showed much stronger preference for a particular candidate over time than the participants who said they would not vote and indeed did not (panel I). Finally, participants who said they would vote but did not had similar profiles as those who said they would vote for a particular candidate and did so (panels C vs. A, and F vs. E), suggesting that most non-voters ended up not voting for reasons unrelated to their voting preferences. Data are for N= 1,263 individuals who participated in USC's survey waves in August, September, early November, and immediately after the election. Shown are unweighted probabilities of voting for Trump and Clinton that participants provided for themselves and their social circles. Participants were classified as being likely to vote for a particular candidate only if they reported more than 50% chance that they would vote in the election; otherwise they were classified as "will not vote". Error bars show ±1 standard error.




**Extended Data Figure 2**. Changes in percentage of like-minded members of their social circle, for participants who voted for different candidates, belonged to different age groups, and were from different states (RR: predicted Republican–voted Republican; DR: predicted Democrat–voted Republican; DD: predicted Democrat–voted Democrat). Error bars are ±1 standard error. Data are for N= 1,263 individuals who participated in all four study waves. Shown are unweighted proportions of participants' social circles that they reported would vote (or had voted) for their preferred candidate.




**Extended Data Figure 3.** Changes in percentage of like-minded members of their social circle, for participants who voted for different candidates, belonged to different education (Edu) groups, and were from different states (RR: predicted Republican–voted Republican; DR: predicted Democrat–voted Republican; DD: predicted Democrat–voted Democrat). Error bars are ±1 standard error. Data are for N= 1,263 individuals who participated in all four study waves. Shown are unweighted proportions of participants' social circles that they reported would vote (or had voted) for their preferred candidate.

**Extended Data Table 1.** Social-circle reports provided a week before the election contribute to the explanation of post-election reports of voting behaviour, over and above own voting intentions and election-winner expectations. Results of linear regressions for voting in the election (vs. not) and voting for different candidates.

| | Voting behaviour reported post-election (November 9-21) | | | |
|---|---|---|---|---|
| | Voted at all b (SE) | Voted for Trump b (SE) | Voted for Clinton b (SE) | Voted for other candidate b (SE) |
| Model with only own voting intentions (October 31-November 7) | | | | |
| Own voting intentions | .94** (.02) | .97** (.01) | .96** (.01) | .82** (.02) |
| Constant | 2.46 (1.97) | 3.11** (.80) | .53 (.86) | -.16 (.53) |
| AIC | 11,351 | 9,501 | 9,636 | 9,165 |
| Model with own voting intentions + social-circle reports (October 31-November 7) | | | | |
| Own voting intentions | .91** (.02) | .89** (.02) | .89** (.02) | .80** (.02) |
| Social-circle reports | .10** (.03) | .15** (.03) | .14** (.03) | .09 (.06) |
| Constant | -3.20 (2.62) | -.74 (1.06) | -2.48* (1.08) | -.49 (.57) |
| AIC | 11,342 | 9,473 | 9,618 | 9,165 |
| Model with own voting intentions + social-circle reports + election-winner expectations (October 31-November 7) | | | | |
| Own voting intentions | | .89** (.02) | .89** (.02) | .80** (.02) |
| Social-circle reports | | .15** (.03) | .14** (.03) | .09 (.06) |
| Election-winner expectations | | .01 (.03) | -.03 (.03) | -.07 (.08) |
| Constant | | -.92 (1.24) | -1.58 (1.51) | -.48 (.57) |
| AIC | | 9,475 | 9,620 | 9,166 |

Note: Analysis of voting vs. not voting is based on N=1,263 individuals who participated in the USC election poll in August, September, late October/early November, and immediately after the election. Analyses of voting for Trump, Clinton, and other candidates are based on those participants who voted in the election (N=1,086). Regression coefficients for own voting intentions, social-circle reports and election-winner expectations refer to the behaviour mentioned in the associated column header. All estimates are unweighted. *p<.05, **p<=.01. Binary correlations (Pearson r) between own intentions and social circles are for voting .41, for Trump .75, for Clinton .74, for other .43. Variance inflation factors for models with social-circle reports are nevertheless low: for voting 1.2, for Trump 2.5, for Clinton 2.5, for Other 1.2. Even though the criteria analysed here are all binary variables, we present linear rather than logit regressions for several reasons[17]: (a) our starting hypothesis was that participants correctly report their actual voting probabilities $P_i$, rather than overestimating small values and underestimating large ones, which would be implied by a logit model with $P_i$ as an explanatory variable; (b) implementing this hypothesis into a logit model with $\ln(P_i/(1-P_i))$ as an explanatory variable would leave many cases undefined as many respondents gave either a 0% or a 100% answer; (c) our explanatory variables are probabilities that bound most of the predictions to a range of approximately 0% to 100%; (d) linear regression provides a convenient way of assessing our predictors (own intentions and social-circle reports). If they are perfect, their linear coefficients should sum to 1, and the intercept should be indistinguishable from 0. The fact that we observe roughly such a pattern in all linear regressions except the one for other candidates suggests that own intentions and social-circle reports are good predictors of own voting intentions. These conclusions were completely supported by additional logit regression analyses. AIC=Akaike Information Criterion.

**Extended Data Table 2a.** Social-circle reports provided in August contribute to the explanation of voting intentions in September, over and above own voting intentions provided in August. Results of linear regressions for intentions to vote in the election and intentions to vote for different candidates.

| | Intentions reported September 12-24 | | | |
|---|---|---|---|---|
| | Intention to vote<br>b (SE) | Intention to vote for Trump vs. another candidate<br>b (SE) | Intention to vote for Clinton vs. another candidate<br>b (SE) | Intention to vote for other candidate vs. Trump or Clinton<br>b (SE) |
| Model with only own voting intentions (August 8-20) | | | | |
| Own voting intentions | .86**<br>(.01) | .94**<br>(.01) | .94**<br>(.01) | .84**<br>(.01) |
| Constant | 13.5**<br>(1.18) | 3.46**<br>(.56) | 2.32**<br>(.60) | 1.78**<br>(.48) |
| AIC | 10,130 | 10,368 | 10,416 | 10,422 |
| Model with own voting intentions + social-circle reports (August 8-20) | | | | |
| Own voting intentions | .83**<br>(.01) | .89**<br>(.02) | .89**<br>(.01) | .80**<br>(.02) |
| Social-circle reports | .07**<br>(0.02) | .10**<br>(.02) | .09**<br>(.02) | .12**<br>(.03) |
| Constant | 9.76**<br>(1.48) | 1.23<br>(.71) | .01<br>(.76) | 1.4<br>(.49) |
| AIC | 10,115 | 10,344 | 10,396 | 10,409 |

Note: Analysis of voting vs. not voting is based on N= 1,263 individuals who participated in the USC election poll in August, September, late October/early November, and immediately after the election. Regression coefficients for own voting intentions and social-circle reports refer to the behaviour mentioned in the associated column header. All estimates are unweighted. *p<.05, **p<=.01. Election-winner expectations do not contribute to explanation of own intentions, except for Clinton (b = .05, SE = .02, p = .04, AIC = 10,393).

**Extended Data Table 2b.** Social-circle reports provided in September contribute to the explanation of voting intentions in late October and November, over and above own voting intentions provided in September. Results of linear regressions for intentions to vote in the election and intentions to vote for different candidates.

| | Intentions reported from October 31 to November 7 | | | |
|---|---|---|---|---|
| | Intention to vote<br>b (SE) | Intention to vote for Trump vs. another candidate<br>b (SE) | Intention to vote for Clinton vs. another candidate<br>b (SE) | Intention to vote for other candidate vs. Trump or Clinton<br>b (SE) |
| Model with only own voting intentions (September 12-24) | | | | |
| Own voting intentions | .84**<br>(.02) | .97**<br>(.01) | .95**<br>(.01) | .78**<br>(.02) |
| Constant | 15.2**<br>(1.52) | 3.53**<br>(.77) | 2.92**<br>(.83) | .44<br>(.61) |
| AIC | 10,673 | 11,135 | 11,279 | 11,018 |
| Model with own voting intentions + social-circle reports (September 12-24) | | | | |
| Own voting intentions | .81**<br>(.02) | .91**<br>(.02) | .89**<br>(.02) | .74**<br>(.02) |
| Social-circle reports | .09**<br>(.02) | .10**<br>(.03) | .13**<br>(.03) | .11**<br>(.04) |
| Constant | 10.94**<br>(1.84) | .95<br>(1.03) | -.37<br>(1.08) | .10<br>(.62) |
| AIC | 10,659 | 11,123 | 11,259 | 11,013 |

Note: Analysis of voting vs. not voting is based on N= 1,263 individuals who participated in the USC election poll in August, September, late October/early November, and immediately after the election. Regression coefficients for own voting intentions and social-circle reports refer to the behaviour mentioned in the associated column header. All estimates are unweighted. *p<.05, **p<=.01. Election-winner expectations do not contribute to explanation of own intentions, except for Clinton (b = .08, SE = .03, p = .02, AIC = 11,256).

**Extended Data Table 3.** Relation between reported voting intentions and actual voting behaviour: Percentage (N) of participants for each combination of intentions and behavior.

| | Voting behaviour reported post-election (November 9-21) | | | | | Total voting intentions |
|---|---|---|---|---|---|---|
| | Did not vote | Voted for Trump | Voted for Clinton | Voted for other | Voted–unknown[a] | |
| Total voting behaviour | 14% | 40% | 38% | 7% | 1% | 100% |
| | (177) | (502) | (484) | (87) | (13) | (1,263) |
| Own voting intentions (October 31–November 7) | | | | | | |
| Will not vote[b] | 87% | 5% | 6% | 1% | 1% | 11% |
| | (125) | (7) | (8) | (2) | (2) | (144) |
| Will vote for Trump | 3% | 94% | 1% | 0% | 1% | 38% |
| | (17) | (458) | (7) | (1) | (3) | (486) |
| Will vote for Clinton | 4% | 3% | 91% | 2% | 1% | 40% |
| | (22) | (13) | (454) | (8) | (3) | (500) |
| Will vote for Other | 11% | 9% | 5% | 72% | 3% | 8% |
| | (10) | (9) | (5) | (68) | (3) | (95) |
| Undecided | 8% | 39% | 26% | 21% | 5% | 3% |
| | (3) | (15) | (10) | (8) | (2) | (38) |

Note: Columns 2–6 sum to 100%. The intended candidate is defined as the one to whom the participant assigned the highest probability of voting for (if several candidates are given the same probability, the participant is coded as undecided). [a]Participants who reported they had voted but not for whom. Results are for N= 1,263 individuals who participated in survey waves in August, September, late October/early November, and immediately after the election. [b]Participants who said there was a less than 50% chance they would vote in the election.

**Extended Data Table 4.** Dynamic interplay between own intentions and social circles over time: results of vector autoregression modeling[23,24]

| | Trump voters | | Clinton voters | | Other voters | |
|---|---|---|---|---|---|---|
| | coef. | (SE) | coef. | (SE) | coef. | (SE) |
| Voting in the election | | | | | | |
| Own intentions | | | | | | |
| Lag 1 own intentions | 0.96** | (0.20) | 1.18** | (0.28) | 0.86** | (0.27) |
| Lag 1 social circle | 0.20** | (0.08) | 0.17* | (0.08) | 0.33 | (0.20) |
| Social circle | | | | | | |
| Lag 1 own intentions | 1.05** | (0.35) | -0.25 | (0.35) | 0.19 | (0.17) |
| Lag 1 social circle | 0.38** | (0.15) | -0.09 | (0.11) | 0.27 | (0.29) |
| Voting for a particular candidate | | | | | | |
| Own intentions | | | | | | |
| Lag 1 own intentions | 1.46** | (0.23) | 2.16** | (0.52) | 4.70 | (3.65) |
| Lag 1 social circle | 0.65** | (0.18) | 1.24* | (0.51) | 0.31 | (0.79) |
| Social circle | | | | | | |
| Lag 1 own intentions | 0.69** | (0.15) | -0.35 | (0.22) | -2.49 | (1.97) |
| Lag 1 social circle | 0.64** | (0.13) | -0.35 | (0.23) | -0.09 | (0.44) |

Note: Analyses are based on unweighted probabilities to vote for Trump and Clinton, which participants provided for themselves and their social circles. See main text for additional details. Results are for N= 1,263 individuals who participated in survey waves in August, September, late October/early November, and immediately after the election. *p<.05. **p≤.01.

**Extended Data Table 5.** Social circles influence voting behaviour of Trump and Clinton voters: Granger causality tests[23,24] for the influence of social circles as reported in the current wave on own intentions to vote in the next wave, and vice versa, for the influence of own intentions to vote in the current wave on social circle reports in the next wave

| | Trump voters | | Clinton voters | | Other voters | |
|---|---|---|---|---|---|---|
| | $\chi^2$ | p | $\chi^2$ | p | $\chi^2$ | p |
| Voting in the election | | | | | | |
| Social circles cause own intentions | 6.24 | .012 | 4.87 | .027 | 2.70 | .100 |
| Own intentions cause social circles | 9.12 | .003 | 0.50 | .481 | 1.20 | .273 |
| Voting for a particular candidate | | | | | | |
| Social circles cause own intentions | 13.15 | .000 | 5.87 | .015 | 0.16 | .693 |
| Own intentions cause social circles | 21.93 | .000 | 0.12 | .116 | 1.60 | .206 |

Note: Analyses are based on unweighted probabilities of voting for Trump and Clinton, which participants provided for themselves and their social circles. Results are for N= 1,263 individuals who participated in survey waves in August, September, late October/early November, and immediately after the election.

**Extended Data Table 6.** Repeated-measures analysis of variance to test the extent and moderators of average levels ("between individuals") and changes over time ("within individuals") in percentage of like-minded voters

| | Trump voters | | Clinton voters | |
|---|---|---|---|---|
| | F | p | F | p |
| Between individuals | | | | |
| Age | **5.76** | **.003** | 1.20 | .302 |
| Education | 0.37 | .693 | 0.32 | .723 |
| State category | 0.58 | .558 | **4.41** | **.013** |
| Age × Education | 1.10 | .356 | 1.49 | .205 |
| Age × State category | 0.80 | .525 | 1.07 | .369 |
| Education × State category | 1.97 | .098 | 1.62 | .168 |
| Within individuals | | | | |
| Wave | **9.82** | **.000** | **4.60** | **.010** |
| Wave × Age | **2.89** | **.021** | 0.17 | .954 |
| Wave × Education | 0.85 | .494 | 0.68 | .604 |
| Wave × State category | **3.72** | **.005** | 0.11 | .979 |
| Wave × Age × Education | 0.92 | .497 | 0.42 | .908 |
| Wave × Age × State category | **2.01** | **.043** | 0.58 | .792 |
| Wave × Education × State category | **1.91** | **.055** | **2.25** | **.022** |
| Adjusted $R^2$ | .64 | | .65 | |

Note: Age is coded as up to 34, 35–64, and 65+ years. Education is coded as high school or less, some college, college or more. States are coded as predicted Republican–voted Republican, predicted Democrat–voted Republican, and predicted Democrat–voted Democrat. Extended Data Figs. 2 and 3 illustrate statistically reliable interactions, noted here in bold. Higher order interactions are not statistically reliable. Greenhouse–Geisser epsilon is .99 for Trump voters and .98 for Clinton voters; Huynh–Feldt correction for sphericity is used for within-individual estimates. Results are for N=1,263 individuals who participated in survey waves in August, September, late October/early November, and immediately after the election. The results can be interpreted as follows (see also Extended Data Figs. 2 and 3): For Trump voters, homogenization of social circles was particularly pronounced among older voters, aged 65 or older, in particular in states that voted Republican (including those "critical states" that were predicted to vote Democrat). In states that voted Democrat, Trump voters in this age group started to homogenize a bit later than in states voting Republican (from September rather than from August 2016). Trump voters in other age groups showed a bit less homogenization, although there was a positive trend among younger groups in strongly Republican states. Education played an additional role, with less educated Trump voters homogenizing more and faster than more educated ones across states. For Clinton voters, age did not predict homogenization, but education and state did. In strongly Republican states, Clinton voters with higher education had the least homogeneous social circles. However, in strongly Democrat states, more educated Clinton voters had somewhat more homogeneous circles than less educated ones, in particular earlier in the election period (in August 2016).