

Brain indices of disagreement with one's social values predict EU referendum voting behavior

Giulia Galli,^{1,*} Miroslav Sirota,² Maurizio Materassi,¹ Francesca Zaninotto,¹ and Philip Terry¹

¹Department of Psychology, School of Social and Behavioural Sciences, Kingston University, Kingston, UK, and ²Department of Psychology, University of Essex, Colchester, UK

Correspondence should be addressed to Giulia Galli, Department of Psychology, School of Law, Social and Behavioural Sciences, Kingston University London, Penrhyn Road, Kingston Upon Thames KT1 2EE, UK. E-mail: g.galli@kingston.ac.uk

Abstract

Pre-electoral surveys typically attempt, and sometimes fail, to predict voting behavior on the basis of explicit measures of agreement or disagreement with a candidate or political position. Here, we assessed whether a specific brain signature of disagreement with one's social values, the event-related potential component N400, could be predictive of voting behavior. We examined this possibility in the context of the EU referendum in the UK. In the 5 weeks preceding the referendum, we recorded the N400 while participants with different vote intentions expressed their agreement or disagreement with pro- and against-EU statements. We showed that the N400 responded to statements incongruent with one's view regarding the EU. Crucially, this effect predicted actual voting behavior in decided as well as undecided voters. The N400 was a better predictor of voting choice than an explicit index of preference based on the behavioral responses. Our findings demonstrate that well-defined patterns of brain activity can forecast future voting behavior.

Key words: event-related potentials; N400; voting behavior; social beliefs

Introduction

At each election or referendum, a main challenge for politicians, campaigners and pollsters is the prediction of voting behavior. This is particularly the case for undecided voters, whose proportion is generally larger than the numerical difference between two leading parties and who therefore play a key role in determining electoral outcomes (Kosmidis and Xezonakis, 2010). Pre-electoral surveys and opinion polls rely on explicit indices of preference for a candidate, party or political position (Mayer, 2008). These surveys typically ask respondents how they will vote in the upcoming elections, along with other explicit measures of political predispositions, perceptions of candidates, agreement or disagreement with public policies and social issues. Forecasts based on these surveys are not always accurate, as recent history suggests (How did the polls get it wrong,

2016). On the other hand, psychological research has demonstrated that more indirect measures of preference, such as implicit attitudes (Arcuri *et al.*, 2008; Friese *et al.*, 2012; Lundberg and Payne, 2014), or trait inference from a candidate's face (Todorov *et al.*, 2005; Ballew and Todorov, 2007; Na *et al.*, 2015), may be effective in predicting voting behavior. In this study, we explored the possibility that a brain signature of disagreement with one's social beliefs, the event-related potentials (ERPs) waveform N400, can predict future voting behavior.

The N400 is a negative-going brain potential with peak at ~400 ms after stimulus onset over posterior scalp locations. It is well established that the N400 is sensitive to semantic violations or incongruences (e.g. "I drink my coffee with milk and socks", Kutas and Hillyard, 1980). More recent work has shown that this brain wave is also sensitive to violations of social knowledge or beliefs. For instance, the N400 is larger when

Received: 30 March 2017; Revised: 21 July 2017; Accepted: 7 September 2017

© The Author (2017). Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

individuals read words associated with social norm violations (as with the word “museum” in the sentence “Samantha is dancing at the museum”, Mu et al., 2015), or words that violate gender or racial stereotypes (the word “aggressive” preceded by the word “female”, White et al., 2009; Hehman et al., 2014; Molinaro et al., 2016). Van Berkum et al. (2009) demonstrated that statements that are not congruent with one’s moral values, such as “I think that euthanasia is an *acceptable* course of action” read by strict Christian believers, elicited an N400 effect. Taken together, these findings support the idea that the N400 is sensitive to disagreements with one’s social beliefs.

Building on these observations, we hypothesized that the N400 would be sensitive to disagreement with politically charged beliefs in an electoral context. This hypothesis was tested in the electoral context of the European Union (EU) referendum that was held in the UK on June 23, 2016. On that day, voters were asked to express their opinion on whether the UK should remain a member of the EU, or leave the EU (“Brexit”). Political science has shown that referenda are concerned with individuals’ opinions on societal issues rather than partisan support of parties and candidates (Laycock, 2013). We thus reasoned that the EU referendum would provide an excellent electoral context to test our hypothesis.

In the 5 weeks preceding the referendum, we recorded electrical brain activity of 62 participants while they expressed their agreement/disagreement with statements consistent with a “remain in the EU” position (e.g. “Free access to healthcare for all EU migrants should be allowed”), or with a “leave the EU” position (e.g. “If Britain leaves Europe our quality of life will be enhanced”). Participants were individuals with a definite intention to vote Remain or Leave, or undecided about their future vote. We expected more negative-going waveforms for Leave statements in participants with an intention to vote Remain, and more negative-going waveforms for Remain statements in participants with an intention to vote Leave, in central-posterior scalp locations, where the N400 is typically prominent. We further tested whether this N400 effect would predict voting behavior. After the referendum, we contacted all participants and asked them to report their actual vote. We found that disagreement with one’s view towards the EU and Brexit elicited larger N400s, and that this N400 effect predicted future voting behavior in both decided and undecided voters. An explicit index of preference based on the rate of agreement with pro- and against-Brexit statements was not equally effective in predicting voting choice. This suggests that brain activity can accurately predict future voting behavior and detect political attitudes even at an embryonic stage, perhaps more accurately than explicit measures of political preferences.

Materials and methods

Participants

Data were collected from 62 participants during the 5 weeks preceding the EU referendum (June 23, 2016). All participants were native English speakers, aged between 18 and 55 years, with normal or corrected-to-normal vision. They all reported being eligible to vote at the referendum. We tested 18 participants who intended to vote “Remain” and 18 subjects who intended to vote “Leave” at the time of their participation, and 26 participants who reported having an intention to vote but were still undecided as to whether to vote Remain or Leave. The three groups differed in age (one-way ANOVA, $F_{2,58} = 5.89$, $P = 0.005$). Participants who reported an intention to vote Leave were older ($M = 33.1$ years,

$SD = 10.7$ years) compared with those who intended to vote Remain ($M = 22.8$ years, $SD = 5.6$ years; $P = 0.004$), but not compared with undecided participants ($M = 29.3$ years, $SD = 9.9$ years, $P = 0.074$). The three groups were similar in terms of years of education (Intention to vote Remain, $M = 15.7$ years, $SD = 1.0$ years; intention to vote Leave $M = 16.0$ years; $SD = 1.57$ years; Undecided $M = 15.8$ years; $SD = 1.30$ years; $F_{2,58} = 0.250$; $P = 0.780$) and gender (number of males in each group: intention to vote Remain 11, intention to vote Leave 11, Undecided 10). One undecided participant did not complete the experiment and therefore was excluded from the analyses. All participants gave their written informed consent before starting the experiment and received course credits or £25 for their participation. The study was approved by the Kingston University Research Ethics Committee.

Materials

Stimuli were 112, 11-word statements, adapted from websites, political articles and campaigns representing both sides of the debate (details on how the statements were pre-tested can be found in the following section). Half of the statements were consistent with a “remain in the EU” position (Remain statements; e.g. “Free access to healthcare for all EU migrants should be allowed”, “If Britain leaves Europe our quality of life will be reduced”). The other half were consistent with a “leave the EU” position (Leave statements). Leave statements were identical to Remain ones except that the last word (critical word) was an antonym and hence attributed the opposite meaning to the statement (e.g. “Free access to healthcare for all EU migrants should be denied”, “If Britain leaves Europe our quality of life will be enhanced”).

Pre-test of the statements used in the experiment

Before data collection, statements were pretested to an independent sample of 34 people (all eligible to vote, 12 males, age $M = 24.6$ years, $SD = 7.2$ years) to ensure that the statements were sufficiently identifiable with either the “remain in the EU” or the “leave the EU” side of the debate. In this pre-test, participants were given a list of 144 statements consisting of 77 Remain statements and their respective Leave opposites, and were requested to judge whether each statement supported “stay in the EU”, “leave the EU” or was neutral. In order to reduce the impact of participants’ personal views, they were encouraged to take a third person’s perspective and to imagine that each statement represented an argument raised in a debate about Brexit among a group of friends. The 112 sentences that we selected for the actual experiment were the ones with the strongest orientation. On average, the Remain statements used in the experiment were judged as supporting “remain in the EU” by 83% of the pre-test participants ($SD 10\%$), and the Leave sentences used in the actual experiment were judged as supporting “leave the EU” by 81% of the pre-test participants ($SD 8.6\%$). 16 additional sentences were selected for practice trials.

Procedure

Statements were presented word-by-word at the centre of the screen. Each word was presented in white Helvetica on a gray background using the Cogent 2000 toolbox (<http://www.vislab.ucl.ac.uk/cogent.php>). Trials began with a 1000-ms fixation mark, then each word appeared for 200 ms followed by a 300-ms blank. This is considered a typical stimulus onset asynchrony in N400 studies (Wlotko and Federmeier, 2015). The last word (critical word) was presented for 1000 ms, and the inter-statement

interval varied randomly between 3500 and 4000 ms. Participants were asked to say whether they agreed with the statement by pressing a key on the keyboard, either the far left key with their left index finger, or the far right key with their right index finger. Response hand indicating agreement or disagreement was counterbalanced across participants. The order of statements was randomized anew for each participant. Sixteen practice trials preceded the beginning of the experiment. Participants were contacted again in the weeks following the referendum day to report their vote.

EEG recording and analysis

EEG was acquired with a 64-channel Biosemi Active Two system (BioSemi, Amsterdam, the Netherlands). Electrodes were positioned according to an equidistant montage using an elasticated cap (http://www.biosemi.com/pics/cap_64_layout_medium.jpg). Scalp electrodes also included a Common Mode Sense active electrode and a Driven Right Leg passive electrode, which are customary used in Biosemi EEG systems to replace the ground electrode. Two additional electrodes were placed on the left and right mastoids, and four electrodes were used to record the electrooculogram (EOG): one above and one below the left eye for vertical EOG, and two on the outer canthus of each eye for horizontal EOG. EEG signals were digitized at 2048 Hz with a bandwidth of DC to 409 Hz using a fifth-order digital sinc filter. Electrode offsets were kept between $\pm 20 \mu\text{V}$. Participants were asked to minimize blinking, movements and muscle contraction during the experiment. Offline analyses were conducted using EEGLAB (Delorme and Makeig, 2004) and ERPLAB (Lopez-Calderon and Luck, 2014). Data were digitally filtered between 0.1 and 30 Hz. The EEG continuous recording was then segmented into epochs time-locked to the onset of the critical word. Epochs lasted from 100 ms prior to the onset of the critical word to 1200 ms afterwards. A 100-ms pre-stimulus baseline was applied. Trials containing blinks, non-blink eye movements, drifts or muscle artifacts were excluded from the averaging process. ERP waveforms were computed by averaging artifact-free epochs surrounding the critical word as a function of statement type and were referenced to the average reference. All averages contained at least 14 trials in the relevant conditions. On average, ERP waveforms for Remain statements were based on 45.5, 37.9 and 45.2 trials and Leave statements on 45.6, 36.3 and 45.8 trials for Remain, Leave and undecided voters, respectively. To analyze N400 effects, mean ERP amplitudes were extracted in the 300–500 ms time window. This window was selected based on previous ERP studies (e.g. Kutas and Hillyard, 1984) and upon inspection of the grand averaged ERPs. Additional analyses were performed on the 500–800 and 800–1200 ms latency intervals to test ERP differences evident in the grand-averaged waveforms that were not relevant to our hypotheses (see Supplementary Text and Supplementary Figure S3). As shown in Supplementary Figure S1, mean ERP amplitudes were extracted for 30 electrodes grouped in 10 regions of interest (ROI): Anterior Left AF7, F5, FC5; Anterior Central Left AF3, F3, FC3; Anterior Central AFz, Fz, FCz; Anterior Central Right AF4, F4, FC4; Anterior Right AF8, F6, FC6; Posterior Left PO7, P5, CP5; Posterior Central Left PO3, P3, CP3; Posterior Central POz, Pz, CPz; Posterior Central Right PO4, P4, CP4; Posterior Central Right PO8, P6, CP6. These electrodes entered all levels of analysis.

ERP analysis approach

First, we aimed to characterize the N400 effect by comparing Remain and Leave statements in all participants with a

definite vote intention at the time of testing (regardless of their actual vote), that is 18 participants who intended to vote Remain and 18 participants who intended to vote Leave. Given that the N400 is larger for socially-incongruent information (Van Berkum et al., 2009; Huang et al., 2014; Mu et al., 2015; Wang et al., 2015; Hundrieser and Stahl, 2016) we expected more negative-going waveforms for Leave statements in participants with an intention to vote Remain, and more negative-going waveforms for Remain statements in participants with an intention to vote Leave, in central-posterior scalp locations, where the N400 is typically prominent (Kutas and Federmeier, 2011). This hypothesis was tested by means of a mixed model Analysis of Variance (ANOVA) with Vote Intention as a between-subjects factor (two levels: Remain, Leave), and the within-subjects factors Statement Type (two levels: Remain, Leave), Anterior–Posterior (two levels: Anterior, Posterior), ROI (five levels: Left, central left, central, central right, Right) and Electrode (three levels, each of the three electrodes comprising the ROI). The Greenhouse–Geisser correction (Greenhouse and Geisser, 1959) was applied if violation of sphericity occurred, and in this case the uncorrected degrees of freedom (df), the corrected *P* values, and the correction factor ϵ are reported. The α level was set to 0.05. We further used Bayes factor (BF) analysis to assess the strength of evidence regarding differences or similarities of the N400 effects. A BF denotes the ratio of the probability of the data given model A (e.g. model assuming H_0 is plausible) to the probability of the data given model B (e.g. model assuming H_1 is plausible). In generic terms, a BF of BF_{01} , with a value of 1 indicates that the probability of data under H_0 is as likely as under H_1 , with a value lower than 1, indicates that the probability of data under H_1 is more likely and a value greater than 1 indicates that the probability of data under H_0 is more likely. These values can be interpreted in terms of evidential categories (e.g. the categorization used here: anecdotal, moderate, strong, very strong and extreme evidence), for instance, BF_{01} values between >1 and 3 or <1 and $1/3$ represent “anecdotal evidence” to support H_0 or H_1 , respectively, whereas BF_{01} values between >100 or $<1/100$ represent “extreme evidence” to support H_0 or H_1 , respectively (Jeffreys, 1961; Lee and Wagenmakers, 2013). We quantified the evidence to support the null or alternative hypothesis by computing a JZS BF ANOVA with default prior scales or default Bayesian *t*-test using the “BayesFactor” package in R or its implementation in JASP (Morey and Rouder, 2015). The prior odds should be updated by means of BF into posterior odds, expressing the probability ratio of the two hypotheses after encountering the data (e.g. if prior odds are uninformative, i.e. equal to 1, and $BF_{10} = 5$, then posterior odds indicate that the alternative hypothesis is five times more likely than the null hypothesis after seeing this data).

The N400 effect used in the prediction analyses was computed by subtracting the mean amplitudes of the N400 for Remain statements from the mean amplitudes of the N400 for Leave statements. We then collapsed the N400 effect across the posterior electrodes that displayed the largest effect according to the ANOVA results and the observed scalp distribution (PO and P electrodes) and performed a logistic regression on these values. We used an unadjusted model with N400 as the only predictor of voting behavior and a model adjusted for decisiveness and their interaction. This analysis was performed on all of the 59 participants who reported having voted on referendum day (two participants did not vote, see Supplementary Text for a description of voting behavior).

Results

Agreement was larger for statements consistent with vote intention in decided voters

Participants with a definite intention to vote Remain agreed more with Remain than with Leave statements ($t_{17}=10.35$, $P < 0.001$, $d = 3.36$, Supplementary Table S1). On the other hand, participants with a definite intention to vote Leave agreed more with Leave than Remain statements ($t_{17}=5.15$, $P < 0.001$, $d = 2.07$), and agreed with Leave statements more quickly than with Remain statements ($t_{17}=2.25$, $P = 0.038$, $d = 0.42$, Supplementary Table S1).

Agreement responses of undecided voters were consistent with actual vote

We next analyzed agreement rates and response times of undecided voters who subsequently voted Remain, and found that participants on average agreed more with Remain compared to Leave statements ($t_{21}=4.06$, $P = 0.01$, $d = 1.53$). Although not as polarized as decided Remain voters (Supplementary Table S2), and despite declaring themselves as undecided at the time of testing, these participants thus showed a preference for statements supporting a “Remain” perspective. No difference emerged in reaction times ($P = 0.469$). Supplementary Table S2 reports response rates and times for undecided voters who subsequently voted Leave. These values did not enter statistical analysis due to the low subject number.

Remainers and Leavers displayed N400 effects of opposite polarity

We first sought to characterize the N400 effect in decided voters (18 Remain and 18 Leave voters). As the N400 is larger for violations of one’s social beliefs (Van Berkum et al., 2009; White et al., 2009; Hehman et al., 2014; Mu et al., 2015; Molinaro et al., 2016), we hypothesized that in participants with a definite intention to vote Remain, statements supporting a Leave perspective would elicit larger negativities over posterior scalp locations than statements indicating a Remain perspective, and that an opposite pattern would arise in participants with a definite intention to vote Leave. The mixed-model Analysis of Variance (ANOVA) on the 300–500 latency window with factors Vote Intention (Remainer, Leaver), Statement Type (Remain, Leave), Anterior–Posterior, ROI and Electrode (see Supplementary Figure S1 for a graphic illustration of electrodes used for the analysis, and their arrangement into ROIs) supported this hypothesis. We found a significant interaction between Vote Intention, Statement Type and Anterior–Posterior ($F_{1,34} = 16.33$, $P < 0.001$, $\eta^2 = 0.324$), and a significant interaction between Vote Intention, Sentence Type, Anterior–Posterior and Electrode ($F_{2,68} = 9.44$, $P = 0.001$, $\epsilon = 0.838$, $\eta^2 = 0.217$). Separate follow-up ANOVAs at anterior and posterior scalp locations in each group revealed that at posterior scalp locations, Remainers displayed larger negativities for Leave compared with Remain statements (main effect of Statement Type $F_{1,17} = 9.46$, $P = 0.007$, $\eta^2 = 0.358$; Figure 1) and Leavers displayed larger negativities for Remain compared with Leave statements in the most posterior electrode position (main effect of Statement Type at PO position $F_{1,17} = 4.81$, $P = 0.042$, $\eta^2 = 0.221$; at the two other electrode positions $P_s > 0.100$; Figure 1. See Supplementary Text for ANOVA results on anterior scalp locations). BF analysis suggested very strong evidence for the main effect of Statement type across all posterior electrodes in Remainers, $BF_{10} = 98$, and moderate

evidence for the main effect of Statement type at PO in Leavers, $BF_{10} = 4.8$ (against the null intercept model). The timing, scalp distribution and response to experimental manipulations of the waveforms are consistent with an N400 effect (Figure 1), therefore our data support the hypothesis that the N400 is sensitive to disagreements with one’s view on politically charged issues in an electoral context.

N400 effects predict voting behavior in decided and undecided voters

Once we had established that our experimental manipulation reliably induced N400 effects in our decided voters, our next step was to assess whether this pattern of brain activity recorded before the vote would predict actual voting behavior reported when participants were contacted again after the referendum. We applied a logistic regression model including the N400 effect, vote decidedness (i.e. whether a participant was decided or undecided at the time of testing) and their interaction. We showed that the N400 effect recorded at the time of testing predicted the actual vote, $OR = 3.01$, 95% CI [1.40, 7.57], $z = 2.60$, $P = 0.009$. The N400 effect still predicted voting to Leave the EU even when we adjusted the model for vote decidedness, $OR = 4.59$, 95% CI [1.70, 18.16], $z = 2.59$, $P = 0.010$. Although decidedness predicted vote to Remain in the EU, $OR = 0.17$, 95% CI [0.03, 0.74], $z = -2.16$, $P = 0.031$, we found no significant interaction between the N400 effect and vote decidedness, $OR = 0.17$, 95% CI [0.02, 1.19], $z = -1.82$, $P = 0.069$. This indicates that predicting actual vote was not dependent on whether participants were decided or undecided at the time of data collection.

We then assessed whether the rate of explicit agreement or disagreement with pro- and against-EU statements could be equally effective in predicting the vote. To this end, for each participant we calculated an index of explicit preference adapted from the Kelley index used in American pre-electoral surveys (Kelley, 1983):

$$\text{Explicit preference index} = \frac{(\% \text{ Leave statements agreed} + \% \text{ Remain statements disagreed}) - (\% \text{ Leave statements disagreed} + \% \text{ Remain statements agreed})}{(\% \text{ Leave statements agreed} + \% \text{ Remain statements disagreed}) + (\% \text{ Leave statements disagreed} + \% \text{ Remain statements agreed})}$$

We entered this explicit index along with the N400 effect into a logistic regression model and found that the explicit index did not significantly predict voting behavior, $OR = 0.96$, 95% CI [0.90, 1.02], $z = -1.24$, $P = 0.215$, whereas the N400 effect did, $OR = 2.86$, 95% CI [1.32, 7.27], $z = 2.45$, $P = 0.015$, supporting the previous analysis. When comparing the nested models, the model featuring both predictors was significantly better than the model featuring the explicit index as a predictor, $\chi^2(1) = 7.42$, $P = 0.006$, whereas it was not significantly better than the model featuring the N400 effect as a predictor, $\chi^2(1) = 1.55$, $P = 0.213$. Therefore, the N400 was a better predictor of voting choice compared with an explicit index of agreement or disagreement with EU-related statements.

We further conducted an additional set of analyses specifically focused on undecided voters. Only three undecided participants voted Leave on referendum day (Supplementary Text reports actual voting behavior in decided and undecided voters), we therefore restricted our examination to undecided participants who subsequently voted to Remain in Europe ($n = 22$). We reasoned that if these participants had an embryonic, implicit preference for voting Remain at the time of testing, their N400 would not differ from that of participants who at the time of testing had a definite intention to vote Remain, and then voted Remain ($n = 17$), but it would differ from the N400 of participants

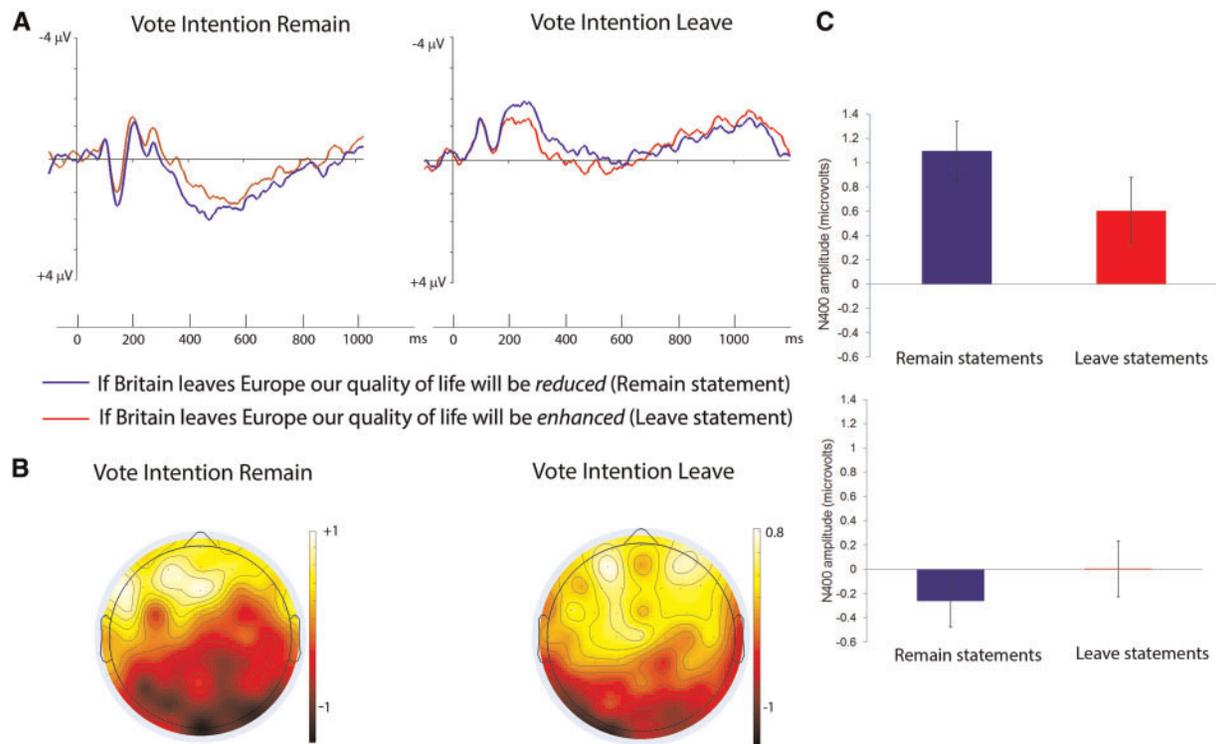


Fig. 1. N400 effects in decided Remain and Leave voters. (A) Waveforms from electrode POz for Remain and Leave statements in voters with a definite intention to vote Remain and a definite intention to vote Leave. Vote-inconsistent statements elicited more negative-going waveforms than vote-consistent statements in both groups. Negative is up. (B) Scalp distribution of the observed ERP difference between incongruent and congruent statements in the 300–500 ms latency region. The N400 effect was prominent at posterior scalp-sites in both groups (C) N400 mean amplitudes collapsed across all analyzed posterior electrode sites in the two groups (Remain top, Leave bottom). Bars depict standard error.

who had the intention to vote Leave, and then voted Leave ($n = 14$). A one-way ANOVA supported this hypothesis. The difference between the three groups was significant ($F_{2,50} = 6.93$, $P = 0.002$, $d = 1.09$). The N400 effect of undecided Remain voters did not statistically differ from the N400 effect of decided Remain voters ($P = 0.940$), but it differed significantly from the N400 effect of decided Leave voters ($t_{34} = 2.62$, $P = 0.013$, $d = 0.89$; Figure 2B and C). A BF analysis provided anecdotal evidence for the absence of difference in the N400 effect between undecided and decided Remain voters, $BF_{01} = 2$ (moderate evidence with wide and ultra-wide priors, Supplementary Figure S2), and moderate evidence for the difference between the N400 effect of undecided Remain and decided Leave voters, $BF_{10} = 4.1$ (moderate evidence with wide and ultrawide priors, Supplementary Figure S2).

Discussion

Consistent with our first hypothesis, we found that statements that disagreed with one's view on the EU and Brexit (e.g. "If Britain stays in the EU our industry will be stronger" for Leave voters, and "If Britain leaves Europe our quality of life will be reduced" for Remain voters) elicited larger N400 waveforms. This result shares similarities with previous work that showed larger N400s for information that is not congruent with personally relevant moral values (Van Berkum *et al.*, 2009), or stereotypical knowledge (White *et al.*, 2009; Hehman *et al.*, 2014; Mu *et al.*, 2015; Molinaro *et al.*, 2016), and further extends this N400 effect to politically charged information in an electoral context. One view on the N400 is that the waveform reflects the process

of semantically integrating the critical word with the preceding context (Hagoort, 2008). According to this view, larger N400s index more extensive cognitive work required to process information that does not fit the semantic context or prior world knowledge. Our results may thus indicate that the cognitive effort required to integrate, e.g. the word "stronger" in the sentence "If Britain stays in the EU our industry will be stronger", will be greater for individuals with against-EU views. Alternatively, N400s for statements that clash with one's view toward the EU and Brexit may indicate difficulty accessing conflicting information stored in semantic memory (White *et al.*, 2009; Kutas and Federmeier, 2011). Either way, our findings unequivocally demonstrate that the processing of politically charged beliefs occurs as early as 400 ms in the brain. This is consistent with the "hot cognition" view that judgments on political issues are fast, automatic and affect-laden (Morris *et al.*, 2003; Lodge and Taber, 2005). Previous work has associated the automated quality of political responses with emotion-related autonomic responses, such as skin conductivity (Oxley *et al.*, 2008), and with brain activity in the amygdala, a brain region implicated in emotion processing (Rule *et al.*, 2010). We speculate that our participants rapidly evaluated incoming information with respect to their views towards the EU, and that this process elicited brain responses well before a conscious evaluation of agreement or disagreement with that information (as suggested by the mean reaction times on the behavioral task, see Supplementary Tables S1 and S2).

Predictors of future vote in decided voters are generally based on explicit attitudes or behaviors recorded with pre-electoral surveys. These surveys measure political predispositions, perceptions

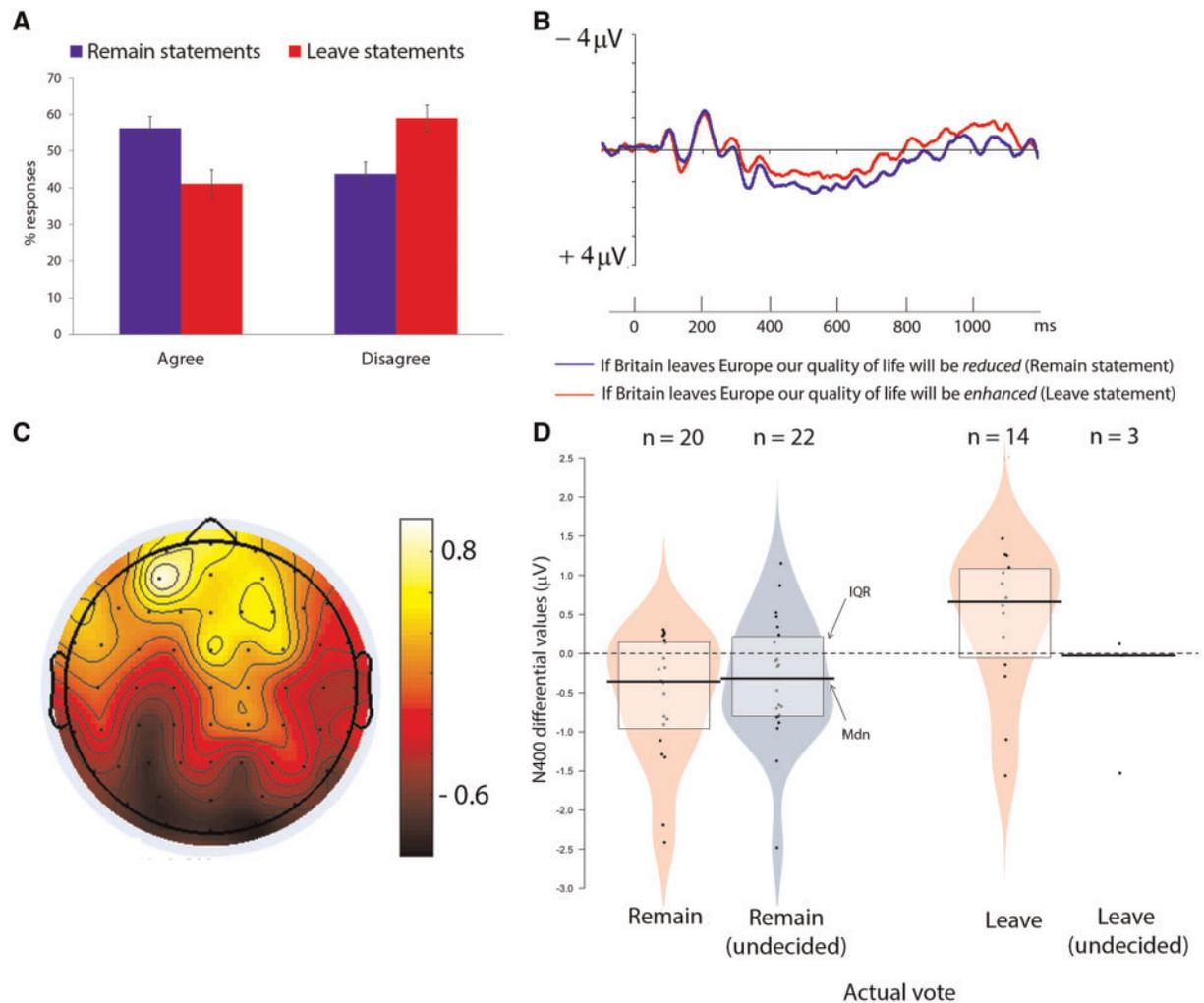


Fig. 2. Behavioral and electrophysiological responses in voters who were undecided at the time of testing and subsequently voted to Remain in Europe. (A) Behavioral performance. (B) Waveforms from electrode POz for Remain and Leave statements. Negative is up. (C) Scalp distribution of the observed ERP difference between incongruent and congruent statements in the 300–500 ms latency region showing a posterior distribution, consistent with an N400 effect and similar to decided Remain voters. (D) Amplitude of N400 effects (μV) for each participant as a function of actual vote and vote decidedness. N400 effects are obtained by subtracting N400 mean amplitudes for Remain statements from N400 mean amplitudes for Leave statements, across the most posterior electrode sites (PO and P electrodes).

of candidates and agreement or disagreement with key political issues such as the economy and safety (Kosmidis and Xezonakis, 2010). Here, we show that explicit agreement or disagreement with statements similar to the ones used in pre-electoral surveys (e.g. “The survey results, 2016”) did not predict voting behavior. Instead, voting choice was predicted by the N400 effect recorded from a few days to several weeks before referendum day. We therefore demonstrate for the first time that brain activity associated with the fast evaluation of politically charged statements can be predictive of the voting behavior of participants who at the time of testing were sure about their vote intention, and crucially, also in participants who were not sure. This latter finding suggests an intriguing possibility—that despite a verbal report of uncertainty, undecided voters have embryonic preferences that guide their electoral choices. Our analyses on the undecided voters showed that the N400 effect of undecided voters who subsequently voted to remain in the EU was not statistically different from the N400 of decided Remainers. This may indicate that embryonic preferences in individuals who had not consciously made up their minds at the time of testing determined a fast, automatic brain response in the presence of disagreements of one’s view

towards the EU, similar to the one of individuals who declare a definite vote intention. This idea is supported by findings from psychological research. For instance, Arcuri *et al.* (2008) found that scores on the Implicit Associations Test (Greenwald *et al.*, 1998) predicted future voting behavior at Italian General and Local elections in decided and undecided voters. This result was replicated in other electoral contexts (Friese *et al.*, 2012; Lundberg and Payne, 2014).

One caveat of this study is that the majority of our undecided participants later voted Remain. This outcome is representative of the voting behavior of UK voters with similar demographics (young adults living in the Greater London area, see for instance <http://news.efinancialcareers.com/uk-en/250125/this-is-how-city-of-london-workers-voted-in>). Due to the small size of the group of undecided voters who subsequently voted Leave, the current data cannot provide conclusive evidence that the N400 predicts voting behavior in this group. Future studies will need to examine voting behavior using a larger sample size in order to allow for different voting outcomes. In addition, it will be of considerable interest to assess whether the current findings are replicated in different

contexts, such as presidential or general elections, and whether the magnitude of the N400 effect is modulated by some political issues more than others, such as immigration or economics.

In summary, our results suggest that brain signatures of disagreement can detect voting preferences well ahead of voting day, even before a conscious endorsement of a position, party or candidate. The measurement of political attitudes by means of explicit agreement or disagreement with a survey statement may be too coarse to detect embryonic preferences in undecided voters. We provide evidence that more refined measurements based on brain activity can be used to predict voting behavior and improve forecasts based on explicit measurements of preferences, which often prove erroneous.

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

- Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology* 29, 369–87.
- Ballew, C.C., Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17948–53.
- Friese, M., Smith, C.T., Plischke, T., Bluemke, M., Nosek, B.A. (2012). Do implicit attitudes predict actual voting behavior particularly for undecided voters? *PLoS One* 7, e44130.
- Delorme, A., Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 134, 9–21.
- Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 12, 1055–69.
- Hehman, E., Volpert, H.I., Simons, R.F. (2014). The N400 as an index of racial stereotype accessibility. *Social Cognitive and Affective Neuroscience* 9, 544–52.
- Kelley, S. (1983) *Interpreting Elections*. Princeton: Princeton University Press.
- Huang, Y., Kendrick, K.M., Yu, R. (2014). Social conflicts elicit an N400-like component. *Neuropsychologia* 65, 211–20.
- Hundrieser, M., Stahl, J. (2016). How attitude strength and information influence moral decision making: evidence from event-related potentials. *Psychophysiology* 53, 678–88.
- Greenhouse, S.W., Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112.
- Greenwald, G., McGhee, D.E., Schwartz, K.J.L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality And Social Psychology* 74, 1464–80.
- How did the polls get it wrong. (2016, November 9). Retrieved from: <http://www.economist.com/blogs/economist-explains/2016/11/economist-explains-3>
- Jeffreys, H. (1961). *Theory of Probability*. Oxford: Oxford University Press.
- Kosmidis, S., Xezonakis, G. (2010). The undecided voters and the economy: campaign heterogeneity in the 2005 British general election. *Electoral Studies* 29, 604–16.
- Kutas, M., Federmeier, K.D. (2011). Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology* 62, 621.
- Kutas, M., Hillyard, S.A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–5.
- Kutas, M., Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307, 161–3.
- Laycock, S. (2013). Is referendum voting distinctive? Evidence from three UK cases. *Electoral Studies* 32, 236–52.
- Lee, M.D., Wagenmakers, E.J. (2013). *Bayesian Modeling for Cognitive Science: A Practical Course*. Germany: Cambridge University Press.
- Lodge, M., Taber, C.S. (2005). The automaticity of affect for political leaders, groups, and issues: an experimental test of the hot cognition hypothesis. *Political Psychology* 26, 455–82.
- Lopez-Calderon, J., Luck, S.J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers In Human Neuroscience* 8, 213.
- Lundberg, K.B., Payne, B.K. (2014). Decisions among the undecided: implicit attitudes predict future voting behavior of undecided voters. *PLoS One* 9, e85680.
- Mayer, W. G. (Ed.). (2008). *The Swing Voter in American Politics*. Brookings Institution Press.
- Molinaro, N., Su, J.J., Carreiras, M. (2016). Stereotypes override grammar: social knowledge in sentence comprehension. *Brain Language* 155, 36–43.
- Morey, R.D., Rouder, J.N. (2015). BayesFactor: an R package for Bayesian data analysis. (Version 0.9.10-2).
- Morris, J.P., Squires, N.K., Taber, C.S., Lodge, M. (2003). Activation of political attitudes: a psychophysiological examination of the hot cognition hypothesis. *Political Psychology* 24, 727–45.
- Mu, Y., Kitayama, S., Han, S., Gelfand, M.J. (2015). How culture gets embrained: cultural differences in event-related potentials of social norm violations. *Proceedings of the National Academy of Sciences of the United States of America* 112, 15348–53.
- Na, J., Kim, S., Oh, H., Choi, I., O'Toole, A. (2015). Competence judgments based on facial appearance are better predictors of American elections than of Korean elections. *Psychological Science* 26, 1107–13.
- Oxley, D.R., Smith, K.B., Alford, J.R., et al. (2008). Political attitudes vary with physiological traits. *Science* 321, 1667–70.
- Rule, N.O., Freeman, J.B., Moran, J.M., Gabrieli, J.D., Adams, R.B.Jr, Ambady, N. (2010). Voting behavior is reflected in amygdala response across cultures. *Social Cognitive and Affective Neuroscience* 5, 349–55.
- The Brexit divide: How the City of London voted in the EU referendum (2016, July 18). Retrieved from <http://news.efinancialcareers.com/uk-en/250125/this-is-how-city-of-london-workers-voted-in>.
- The survey results. (2016, March 20). Retrieved from <https://www.theguardian.com/politics/2016/mar/20/britons-on-euro-pe-survey-results-opinium-poll-referendum>.
- Todorov, A., Mandisodza, A.N., Goren, A., Hall, C.C. (2005). Inferences of competence from faces predict election outcomes. *Science* 308, 1623–6.
- Van Berkum, J.J., Holleman, B., Nieuwland, M., Otten, M., Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science* 20, 1092–9.
- Wang, L., Bastiaansen, M., Yang, Y. (2015). The influence of emotional salience on the integration of person names into context. *Brain Research* 1609, 82–92.

White, K.R., Crites, S.L., Taylor, J.H., Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience* 4, 191–8.

Wlotko, E.W., Federmeier, K.D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex* 68, 20–32.