

What Went Viral on Britain First’s Facebook Page? Outrage, Questions and the Cost of Mobilisation, 2013–2018

Tristan Hotham*

2026

Preprint, not peer reviewed

Abstract

Britain First built a large following on Facebook before the platform removed its pages on 14 March 2018. The page and its post-level record are no longer publicly available. I analyse an archive of 3,505 posts published between 25 October 2013 and 12 March 2018, extracted two days before the ban, which together attracted 32.8 million reactions, comments and shares. I code the text of every post for content, interaction prompts, leadership personalisation and sentiment, then estimate what predicts engagement using month fixed effects, cluster-robust inference, text-as-data methods and a matched comparison with the Labour, Conservative, UKIP, Liberal Democrat and Green pages. Engagement is extremely concentrated: the top 1% of posts attracted 62% of all interactions. Posts about Islam, immigration and crime drew 56%, 39% and 20% more engagement than posts published in the same month, and posts that asked a question (137%) or asked readers to share (116%) drew the most. Posts asking supporters to donate or join (−58%) or linking off Facebook (−51%) drew the least. Naming the group’s own leaders made no measurable difference. On the average Britain First post after February 2016, 28% of reactions were angry; on mainstream party posts in the same months and formats the figure was 2 to 6%. Britain First’s audience shared more and commented less than the Conservative and Labour audiences. The audience engaged most with content about out-groups and with direct invitations to react, and least with requests to join, give or click through. I discuss what this means for how platforms rank extremist content and for the study of far-right communication when the underlying data have been deleted.

Keywords: Britain First; far right; Facebook; engagement; outrage; populist radical right; social media; text as data

1 Introduction

On 14 March 2018 Facebook removed the pages of Britain First and of its leaders Paul Golding and Jayda Fransen for repeatedly breaching its community standards (ITV News 2018). The decision followed a year in which the group had moved from the margins of British politics to the centre of an international row, after the President of the United States re-posted three videos from its deputy leader to his followers on 29 November 2017 (ITV News 2017). Britain First had by then built a very large Facebook following while winning no electoral success of note. Facebook was central to the group’s public presence.

*Independent researcher. Data collected during doctoral research at the University of Bath. Correspondence: contact@tristanhotham.com. The views expressed are the author’s own and do not represent those of any employer.

The ban removed a case that political scientists had barely begun to study. The page, its posts and the interactions beneath them vanished from public view, and platform data access for researchers has since narrowed further (Freelon 2018). A small number of archived extractions survive. In this paper I analyse one of them: the post-level record of the Britain First page, collected in March 2018 during my doctoral research on party use of Facebook.

The question is a simple one with large consequences. What did the audience of a far-right page respond to? Studies of political communication on social media show that engagement is not random. Format, emotion, interactive prompts and the targeting of out-groups all shape what is shared (Berger and Milkman 2012; Brady et al. 2017; Rathje et al. 2021; Bene 2017). Studies of the populist and extreme right show that these actors communicate in distinctive, negative and out-group-focused ways (Engesser et al. 2017; Ernst et al. 2017), and that social media suit that style (Gerbaudo 2018). What is missing is a fine-grained, post-level account of how a single prominent far-right organisation was rewarded by its audience over four years, and of how that reward structure compares with the mainstream parties competing for the same platform.

I make four contributions. First, I document the extreme concentration of engagement on the page and show that a few viral posts account for most of its measured reach. Second, I estimate the association between textual content and engagement within months, separating timing and audience growth from the content of the message. Third, I show a mobilisation trade-off: the content that asked for organisational commitment (petitions, membership, donations, links off Facebook) drew the least engagement, while content about out-groups drew the most. Fourth, I place the page in comparative context, using a matched design on month and format to compare how its audience interacted with the audiences of the Labour, Conservative, UKIP, Liberal Democrat and Green pages.

2 Theory and Hypotheses

2.1 Engagement as the currency of platform politics

On Facebook an organisation's visibility depends on the reaction its posts generate. Engagement determines how widely a post travels beyond existing followers, so organisations learn over time to produce what is rewarded. Engagement is also the most visible signal of audience response, although administrators also see reach and clicks in the platform's page analytics. The consequence is a feedback loop between audience behaviour and message selection, and the stable features of what is rewarded tell us what a page's strategy is likely to become.

2.2 Hypotheses

Format. Video and photo posts carry richer cues than links and are favoured by the platform. I expect video to outperform photo and photo to outperform links (H1).

Out-group content. Moral and emotional language predicts the diffusion of political content (Brady et al. 2017), and out-group animosity was the strongest predictor of sharing among the features tested by Rathje et al. (2021), and high-arousal negative emotions such as anger increase virality (Berger and Milkman 2012). The populist radical right defines itself against cultural and ethnic out-groups (Mudde 2007). I expect posts about Islam, immigration and crime to receive more engagement than other posts (H2), and to provoke a larger share of angry reactions (H3).

Personalisation. The personalisation thesis expects leaders to be an engagement asset. A small organisation led by two highly visible figures offers a hard test. I expect posts that name the group's own leaders to receive more engagement (H4). This corresponds to the leader-usage code I applied to party pages in my thesis.

Interaction prompts. Direct address and requests for interaction are standard techniques for increasing engagement (Bene 2017). I expect posts that ask the audience a question or ask it to share to receive more engagement (H5).

Mobilisation trade-off. Posts that ask supporters to do something costly, such as donate, join, sign or leave Facebook for the organisation's own website, impose a higher cost on the reader than a reaction. I expect these posts to receive less engagement than posts that ask for nothing (H6).

Events. Terrorist attacks and political crises supply material for a group focused on Islam and immigration. I expect engagement to rise in the week after salient events (H7).

Frequency. Posting more often can dilute the engagement each post receives. I expect per-post engagement to fall on days with more posts (H8).

2.3 Research question

How does the interaction profile of Britain First's audience differ from those of the mainstream parties' audiences when format and month are held constant (RQ1)? I examine the composition of engagement into reactions, comments and shares, and the composition of reactions by type.

3 Data

3.1 The archive

The data are the full post-level record that the extraction tool Netvizz (Rieder 2013) returned for the Britain First Facebook page on 12 March 2018. Netvizz returns, for each post, its type, publication time, message text, link, and counts of likes, reactions by type, comments and shares. The record contains 3,505 posts from 25 October 2013 to 12 March 2018: 1,567 photos, 1,891 videos, 38 links and 9 events and status updates. The extraction took place two days before the ban. The page no longer exists, the posts can no longer be retrieved from the platform, and I am not aware of any other post-level archive of the page.

Engagement is the sum of reactions, comments and shares, as defined by the extraction tool. Together the posts attracted 32,809,525 interactions, of which 59% were shares, 34% reactions and 7% comments. Facebook introduced its differentiated reactions (love, haha, wow, sad, angry) on 24 February 2016, so analyses of reaction types use the 1,783 posts published after that date.

3.2 Coverage and what it means for inference

Facebook's interface limits how much of a page's history an extraction tool can reach, and in my thesis I documented that Netvizz was constrained to roughly 600 posts per page per year. The Britain First archive is consistent with that constraint: it holds 560 to 605 posts in every calendar year. The page's total output is unknown, but the match to the cap indicates that the archive is a retrieval-limited subset

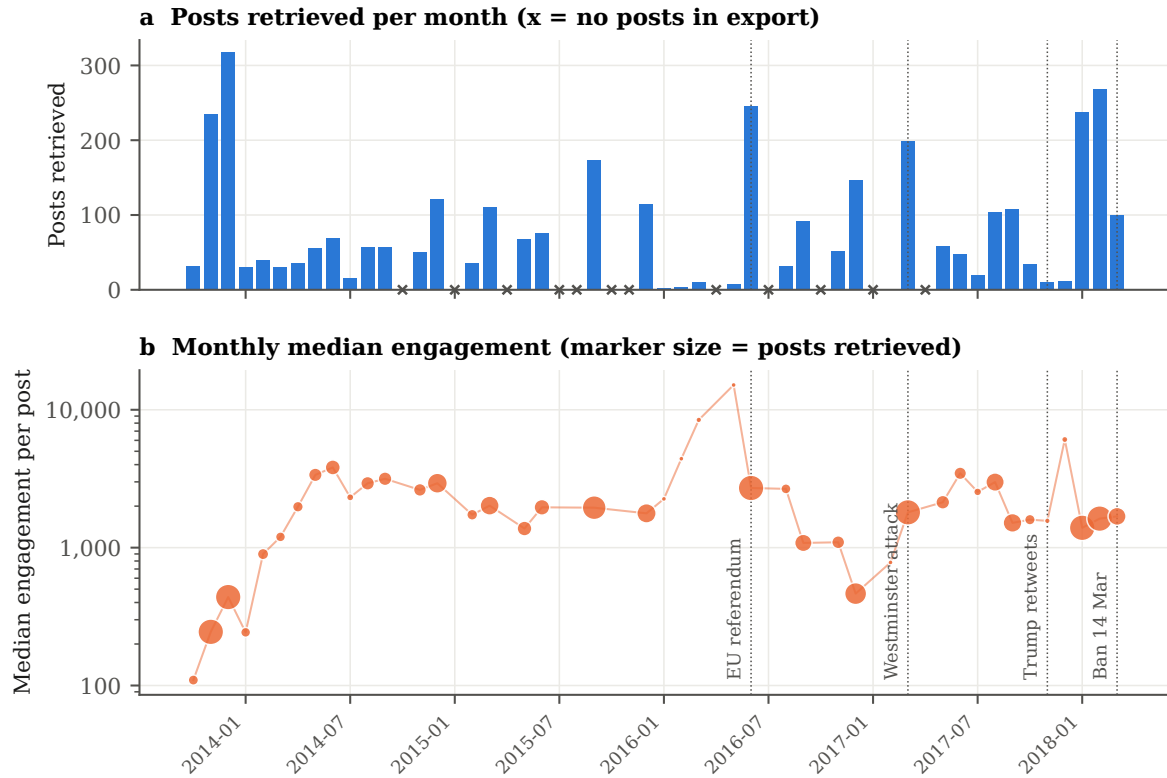


Figure 1. Coverage and engagement over time. Panel a shows posts retrieved per month. Panel b shows the monthly median engagement per post on a log scale. Dotted lines mark the EU referendum (June 2016), the Westminster attack (March 2017), the Trump re-posts (November 2017) and the ban (March 2018).

of the page's output, not a census. Posts are present in 42 of the 54 calendar months in the period, and the number retrieved per month ranges from 1 to 318 (Figure 1). The retrieval rule is not documented, and I understood it in my thesis to favour posts the platform ranks as most engaged. If so, the archive over-represents high-engagement posts and truncates the lower tail. The retrieved posts also fall in contiguous runs of days separated by gaps of up to two months, which is consistent with retrieval by date as well.

Three features of the analysis respond to this. First, all models include month fixed effects, so that identification comes from comparisons among posts published in the same month and the uneven coverage of months does not drive the results. Second, I treat the levels reported here (medians, means, concentration) as descriptions of the archive, not of the page's full output. Third, I test directly how selection from below would affect my estimates by progressively trimming the lowest-engagement posts in each year and re-estimating (Appendix A.1). Truncation on engagement attenuates associations, so trimming further shrinks them whether or not the archive was already truncated; the exercise shows the direction of the bias, not whether selection took place. The hostile-theme, question and share-request associations shrink as the lower tail is removed, the URL and link penalties shrink toward zero, and the fundraising estimate changes sign. If the archive was selected on engagement within years, the estimates I report for the hostile themes and prompts are likely to understate the associations in the page's full output. If it was selected by date, month fixed effects absorb the selection.

3.3 Comparison pages

For the comparative analysis I use archives of the Labour, Conservative, UKIP, Liberal Democrat and Green party pages, extracted with the same tool in early March 2018 and subject to the same retrieval constraint. I restrict all pages to the Britain First study window and to photo and video posts, which make up 98.7% of the Britain First archive and are the formats directly comparable across pages. Because every page is retrieval-limited and uneven in monthly coverage, I do not compare engagement levels, which depend on audience size and on what each archive retains. I compare the composition of engagement, which is scale-free, within cells defined by month and format.

3.4 Ethics and data handling

All data are public posts by an organisation, not by private individuals. The analysis uses aggregate counts and coded categories. I do not reproduce abusive content, and where I quote, I quote words as variables. The coded dataset without message text can be shared on request.

4 Measures and Methods

4.1 Outcomes

The main outcome is $\log_{10}$ engagement per post. I also analyse $\log_{10}(1+\text{count})$ of shares, comments and reactions separately, since they measure different things: reactions are low-cost endorsement, shares are amplification, and comments are conversation. For reactions I use the share of reactions that are angry, haha, love, sad or wow.

4.2 Text coding

I code every post with transparent dictionaries and rules, built from the content categories I used for party pages in my thesis and adapted to the content of this page (Table 1). Image content is not coded, because the archive contains message text and not the images. Leader usage is therefore measured as the appearance of a leader's name in the message. The categories are not mutually exclusive.

4.3 Statistical strategy

The core model is a linear regression of $\log_{10}$ engagement on the content codes, post format, message length, time of day and weekend, with a fixed effect for every calendar month and standard errors clustered by month ($G = 42$). Month fixed effects absorb audience growth, algorithm changes and the uneven monthly sampling. I report coefficients as percentage differences in engagement, $100(10^{\hat{\beta}} - 1)$. For the 26 substantive coefficients (content, format and the auto-generated flag) I control the false discovery rate with the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995).

I supplement the model with: Kruskal–Wallis and Mann–Whitney tests with Cliff's delta (Cliff 1993) and Holm correction (Holm 1979) for unadjusted comparisons; a Poisson pseudo-maximum-likelihood model on raw counts (Santos Silva and Tenreiro 2006); quantile regression at the median and 90th percentile (Koenker and Bassett 1978); a Hill estimator and Gini coefficient for concentration (Hill 1975; Clauset et al. 2009); a weighted log-odds analysis of words with an informative prior (Monroe et al.

Table 1. Text coding scheme. Dictionary terms are case-insensitive and abbreviated; asterisks mark word stems. Prevalence is the percentage of the 3,505 posts coded 1. Additional codes: exclamation mark, URL in text, “Britain First” sign-off, hashtag, message length, and auto-generated Facebook text (for example “Britain First updated their cover photo”).

Code	Terms and rules	% of posts
Own leaders named	Paul Golding, Jayda Fransen, Jim Dowson, “our leader”, “deputy leader”, “Britain First leader”	7.4
Other politicians named	Cameron, May, Corbyn, Miliband, Farage, Khan, Merkel, Obama, Trump and other named public figures	2.8
Islam / terrorism	Islam*, Muslim*, mosque*, jihad*, sharia, halal, burka*, terror*, extremis*, ISIS	38.0
Immigration / borders	immigra*, migrant*, asylum, refugee*, border*, deport*, illegal*, Calais	10.2
Crime / justice	rape*, grooming, murder*, stabb*, attack*, abuse*, police, court, jail*, crime*, victim*	17.5
Military / veterans	troops, soldier*, veteran*, army, armed forces, hero*, remembrance, poppy	5.8
Animal welfare	dog*, cat*, animal*, puppy, horse*, cruelty, RSPCA	1.2
National symbols	Queen, royal*, Union Jack, flag, St George, English, British, Britain (which also matches the page’s own name), Christian*, church	41.8
EU / Brexit	EU, Brussels, Brexit, referendum, leave, remain, UKIP	1.9
Establishment / media	media, BBC, Labour, Tories, politicians, government, elite*, traitor*, liberal*	5.5
Activism / patrols	patrol*, campaign*, march, rally, leaflet*, activist*, volunteer*, join us, members	14.8
Fundraising / membership	donat*, fund*, support us, merchandise, join now, subscri*, pledge	3.3
Persecution / legal	jailed, trial, charged, persecut*, free speech, censor*, banned	4.9
Women, children, vulnerable	women, girls, children, mothers, disabled, pensioners, homeless, NHS	6.9
Question	any question mark, “do you agree”, “will you”, “are you”, “what do you think”	6.6
Asks audience to share	share, shares, sharing	13.5
Like (word or request)	like, “if you agree/care/love/support” (includes “like” used as a preposition)	25.2
Imperative verb	sign, join, donate, vote, demand, stop, ban, end, support, help, stand	16.5
Shouting	three or more fully capitalised words	39.5
Negative sentiment	VADER compound score of -0.25 or lower	29.3

2008), with a permutation null on the maximum $|z|$ to control family-wise error; sentiment from VADER (Hutto and Gilbert 2014); and a month-blocked cross-validated ridge regression to assess out-of-sample predictability.

For the event analysis I compare engagement in the seven days from an event with the surrounding 60 days, using a day-level permutation test. For the comparison with other parties I estimate, within month-by-format cells that contain posts from both pages, the difference in the mean share of engagement (or of reactions) that is of a given type, weighted by the smaller cell. I test it by permuting page labels within cells and build confidence intervals with a month-cluster bootstrap. All tests are two-sided.

5 Results

5.1 Engagement is extremely concentrated

Median engagement per post is 1,473 and the mean is 9,361. The distribution has a very heavy right tail (Figure 2). The top 1% of posts account for 62% of all engagement and the top 10% for 81%. The Gini

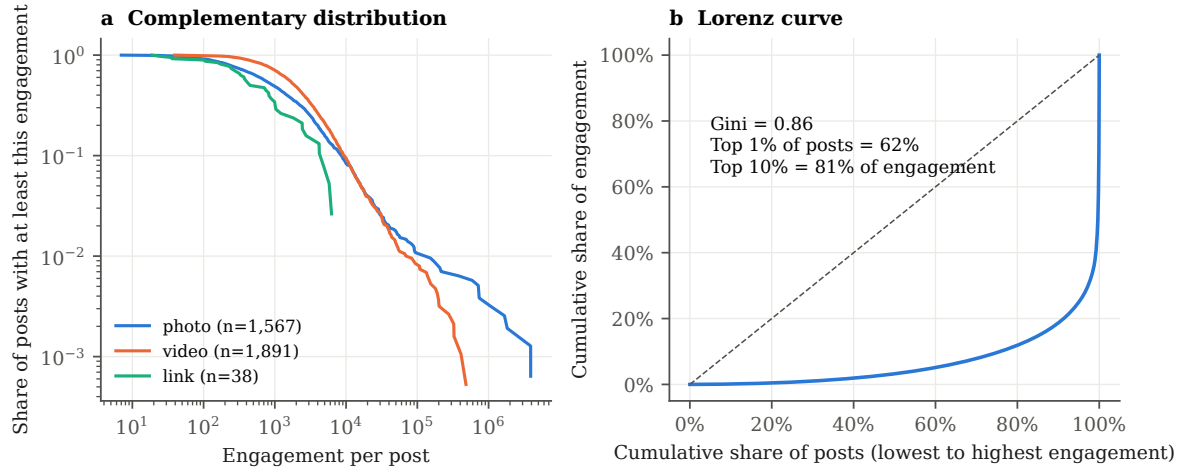


Figure 2. Concentration of engagement. Panel a shows the complementary distribution of engagement by format on log-log axes. Panel b shows the Lorenz curve.

Table 2. Engagement by post format. Medians with bootstrap 95% confidence intervals (1,000 resamples). Events and status updates ($n = 9$) are omitted.

	n	Median	95% CI	Mean	Median reactions	Median comments	Median shares
Photo	1,567	953	[848, 1,060]	13,866	575	70	189
Video	1,891	1,919	[1,771, 2,053]	5,830	891	352	520
Link	38	434	[302, 907]	1,249	312	92	36

coefficient is 0.86, and a Hill estimate for the tail exponent is 2.07. At an exponent near 2 the variance of engagement is not finite and a handful of posts dominate the mean. Five posts alone account for 38% of all interactions in the archive. Two of them, both posted within the first seven months of the archive, carry over 3.8 million interactions each. Because of this concentration, I analyse log engagement throughout and test the sensitivity of every result to outliers.

5.2 Format: links do worst, and video shifts engagement towards shares and comments

Engagement differs sharply by format (Kruskal–Wallis $H = 192.0$, $p < .001$, $\epsilon^2 = 0.054$; Table 2, Figure 3). Videos have a median of 1,919 interactions against 953 for photos and 434 for links. All pairwise differences are significant after Holm correction (video over photo, $\delta = 0.26$; video over link, $\delta = 0.50$; photo over link, $\delta = 0.21$). The raw ordering supports H1.

The adjusted picture is more specific. Holding month and content constant, video is associated with no difference in total engagement relative to photos (+1%), but with more shares (+64%) and comments (+50%) and fewer reactions (−24%) (Table 3). Videos change the composition of what the audience does. They are less likely to be liked and more likely to be shared and discussed, and comments make up 16% of video engagement against 3% for photos (Figure 3b). Links are the weakest format, with 53% lower engagement after adjustment. H1 is supported for links, and not for video over photo in total engagement.

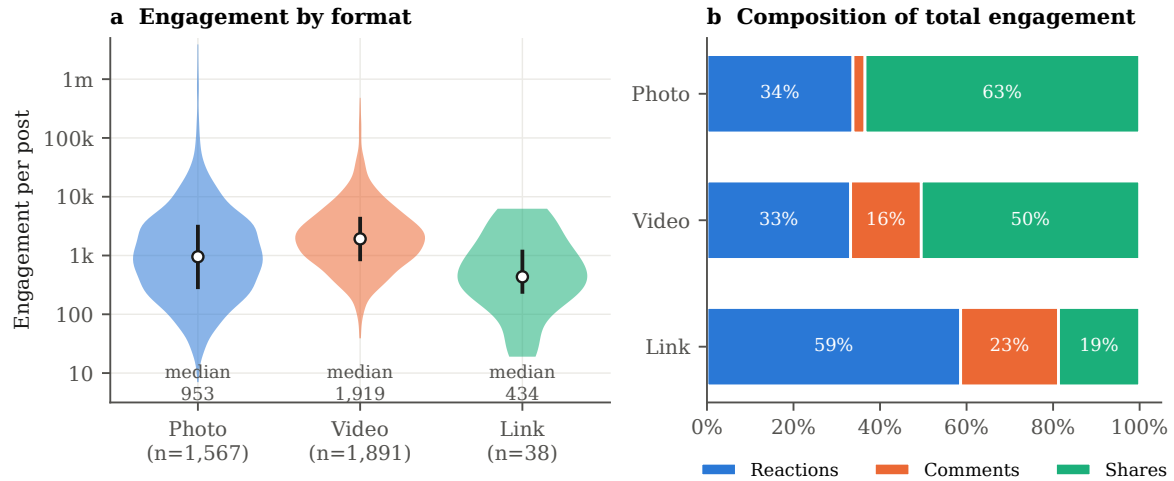


Figure 3. Engagement by post format. Panel a shows the distribution of log engagement with the median and interquartile range. Panel b shows the composition of total engagement.

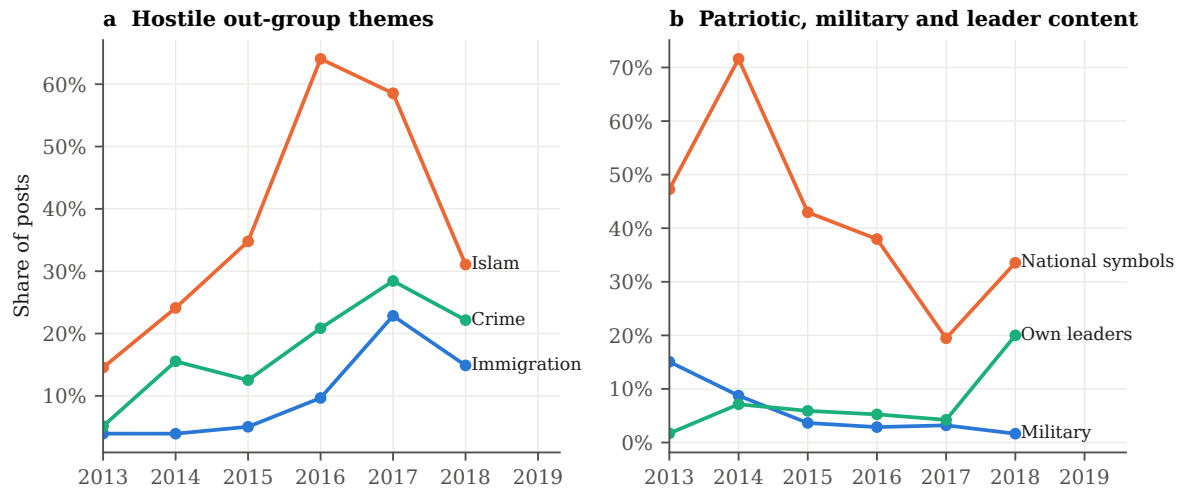


Figure 4. Content over time. Share of posts in each year coded for each theme.

5.3 Content: posts about out-groups drew more engagement

The content of the page changed (Figure 4). The share of posts about Islam rose from 15% in 2013 to 64% in 2016 (chi-squared = 473, $df = 5$, $p < .001$), and posts about immigration rose from 4% in 2013 to 23% in 2017. Military and veterans content, which made up 15% of posts in 2013, nearly disappeared. Over the same years the page became more hostile in tone: the share of posts with negative VADER sentiment rose from 9% to 42%.

Table 3 and Figure 5 give the adjusted associations. Posts about Islam receive 56% more engagement than other posts in the same month (95% CI +43 to +71), posts about immigration 39% more (+22 to +58) and posts about crime and justice 20% more (+6 to +35). Posts about women, children and vulnerable people receive 38% more (+14 to +67) and posts about animal welfare 78% more (+31 to +142). Most of these posts also mention Islam or crime (76% of the women-and-children posts and 68% of the animal-welfare posts), so these codes largely capture victim framing within the out-group themes and not separate consensual causes. Negative sentiment adds a smaller 14% (+3 to +26). All of these

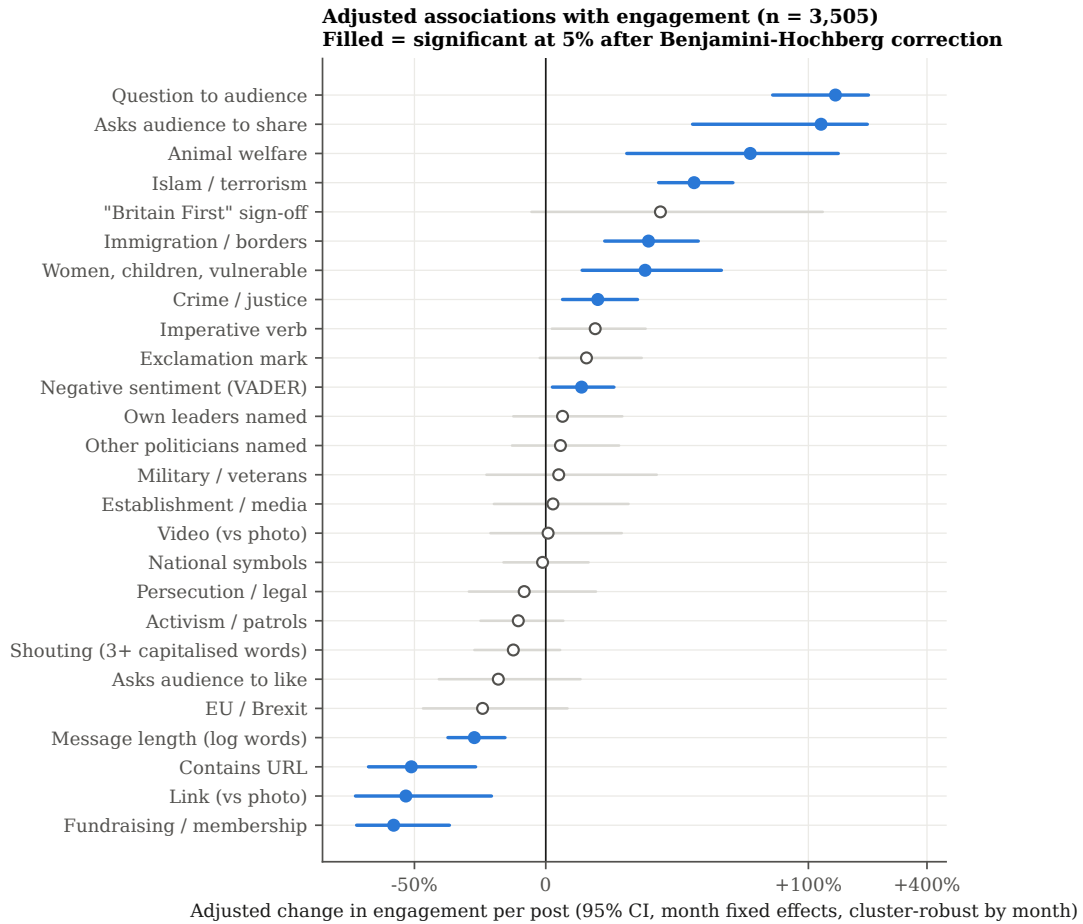


Figure 5. Adjusted associations between content features and engagement. Estimates are percentage differences from a regression of log engagement on all features shown plus format, time of day, weekend and month fixed effects, with 95% confidence intervals clustered by month.

survive Benjamini–Hochberg correction at 5%. H2 is supported. The Islam association is the largest of the three and the most robust: it ranges from 56% to 75% across the seven specifications I estimate (Table 5) and holds at the median (56%) and the 90th percentile (46%) in quantile regression.

5.4 Interaction prompts: questions and share requests

The strongest associations belong to prompts that give the reader something to do with the post. Posts that contain a question (“Do you agree?”, or any question mark) receive 137% more engagement (+86 to +202), and posts that ask the audience to share receive 116% more (+56 to +199). Share requests are associated with 265% more shares (+152 to +427), the largest association with any outcome in the study, and questions with 216% more comments (+143 to +310). The like code, which also picks up “like” used as a preposition, shows no adjusted association (–18%, –41 to +13). H5 is supported for questions and share requests.

The raw data point the other way. The median share-request post receives only about a third of the engagement of the median post that does not ask, because share requests clustered in the opening weeks of the page, when its audience was small. Month fixed effects remove this compositional artefact, and the positive effect persists when I drop the first two months (81%). Figure 6 shows how features

Table 3. Adjusted associations between post features and engagement. Entries are percentage differences from OLS regressions of $\log_{10}$ engagement (column 1), reactions, shares and comments on all features shown plus time-of-day, weekend, auto-generated post and month fixed effects ($n = 3,505$). Column 1 shows 95% confidence intervals from standard errors clustered by month and the Benjamini–Hochberg adjusted p -value for the 26 substantive coefficients. * $p < .05$, ** $p < .01$, *** $p < .001$ (unadjusted). R^2 : total engagement 0.41, reactions 0.42, shares 0.39, comments 0.51.

Feature	Total engagement		Reactions	Shares	Comments
	% [95% CI]	q (BH)	%	%	%
Islam / terrorism	+56 [+43, +71]	<.001	+34***	+91***	+92***
Immigration / borders	+39 [+22, +58]	<.001	+34***	+62***	+35***
Crime / justice	+20 [+6, +35]	.008	+18**	+24**	+27**
Women, children, vulnerable	+38 [+14, +67]	.003	+21*	+69***	+30**
Animal welfare	+78 [+31, +142]	.001	+47**	+129***	+65***
Negative sentiment (VADER)	+14 [+3, +26]	.035	+11*	+23***	+18*
Military / veterans	+5 [-23, +42]	.855	+11	+5	-15
National symbols	-1 [-16, +16]	.919	+6	-10	-5
EU / Brexit	-24 [-47, +8]	.223	-21	-28	-11
Establishment / media	+3 [-20, +32]	.900	-5	+7	+14
Activism / patrols	-10 [-25, +7]	.329	-4	-16	-20*
Persecution / legal	-8 [-29, +19]	.658	+1	-13	-18
Fundraising / membership	-58 [-72, -37]	<.001	-57***	-62***	-47**
Own leaders named	+6 [-12, +29]	.658	+15	-12	+7
Other politicians named	+6 [-13, +28]	.683	+5	-14	+28*
Shouting (3+ capitalised words)	-12 [-27, +5]	.265	-7	-23*	-13
Exclamation mark	+16 [-2, +37]	.170	+16	+25*	+11
Question	+137 [+86, +202]	<.001	+120***	+118***	+216***
Asks audience to share	+116 [+56, +199]	<.001	+66***	+265***	+95**
Like (word or request)	-18 [-41, +13]	.329	-11	-21	-34*
Imperative verb	+19 [+2, +38]	.052	+13	+21	+14
“Britain First” sign-off	+44 [-5, +118]	.170	+28	+81*	+17
Contains URL	-51 [-67, -27]	.002	-51***	-46*	-54***
Video (vs photo)	+1 [-21, +29]	.944	-24*	+64**	+50**
Link (vs photo)	-53 [-72, -21]	.013	-54***	-70**	-25
Message length (log words)	-27 [-37, -16]	–	-29***	-28***	-5

map onto the three kinds of engagement. Content about Islam is associated about equally with shares and comments (+91% and +92%), and content about immigration more with shares (+62%) than with comments (+35%), while questions are associated most with comments, and share requests with shares.

5.5 The mobilisation trade-off

Content that asked supporters to build the organisation drew the least engagement on the page. Posts about fundraising and membership receive 58% less engagement (-72 to -37), posts that contain a URL receive 51% less (-67 to -27) and posts in the link format 53% less. The word analysis shows the same pattern from the bottom up (Figure 7). Among the words over-represented in low-engagement posts, I find *sign*, *join*, *bulletins*, *action*, *click*, *petition* and *membership*. Among those over-represented in high-engagement posts, I find *share*, *muslim*, *muslims*, *think* and *agree*, but also *website* and *activists*, with *migrants* just below the threshold. Of the 235 words tested, 17 exceed the permutation threshold for

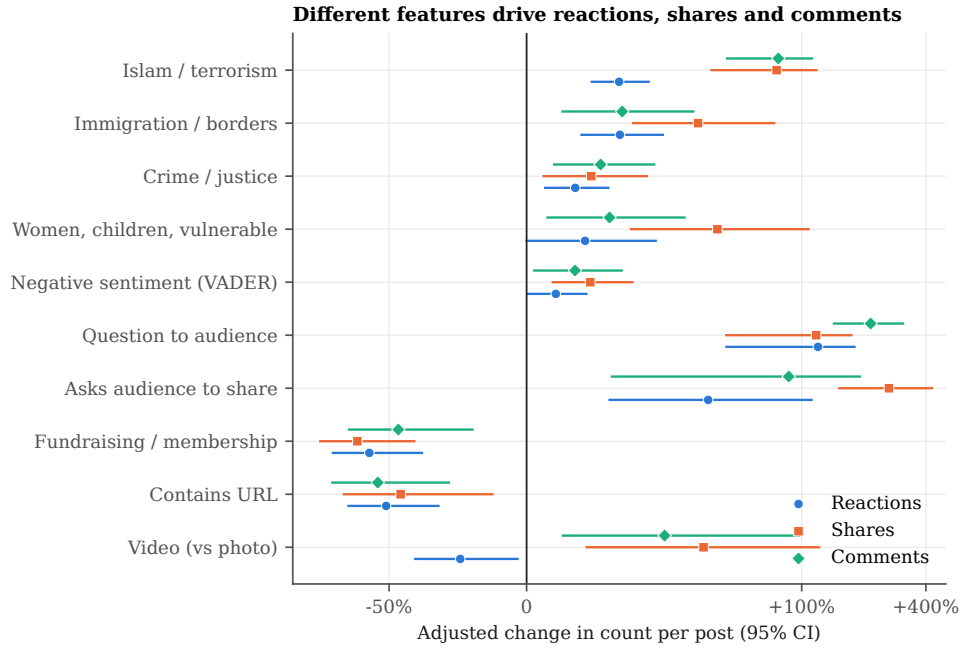


Figure 6. Adjusted associations by type of engagement. Estimates and 95% confidence intervals from separate regressions of log reactions, log shares and log comments on the same specification as Figure 5.

family-wise error control ($|z| > 4.20$). The word analysis uses the same posts as the regression but does not depend on the dictionary codes, and it mostly points the same way, with the exception that *website* appears among the high-engagement words.

The fundraising estimate is the most fragile result in the paper. It rests on only 115 posts, most of them from 2013 to 2015. It is no longer significant when the lowest-engagement tenth of posts is removed and changes sign when a quarter is removed (Appendix Table 6); the penalties for URLs and links shrink under the same trimming. I therefore read the mobilisation penalty as a pattern across the URL, link, fundraising and word-level evidence, and not as a precise estimate for any one of them.

The pattern is consistent with H6. An organisation seeking engagement and an organisation seeking members may face different optimal messages on a platform that ranks content by attention. Britain First’s audience engaged most with content about out-groups and least with what it was asked to do. I cannot observe whether the mobilisation posts achieved their purpose, such as clicks, sign-ups or donations, so engagement is the wrong outcome for judging them.

5.6 Personalisation: naming the leaders made no difference

Posts that name the group’s own leaders receive 6% more engagement (95% CI -12 to $+29$, $p = .53$), which is indistinguishable from zero but does not rule out a moderate increase. The estimate is stable across the specifications with month fixed effects, and is $+26\%$ without them (Table 5). H4 is not supported. Leadership content was uncommon until January to March 2018, when it rose to 20% of posts as the leaders faced prosecution and prison, and posts naming the leaders attracted a much lower share of angry reactions (12% against 30%). I find no evidence that naming the leaders added to engagement. This differs from the expectation of the personalisation thesis. The result needs the usual caveat that I code leadership from names in the text and cannot observe the leaders’ faces in images.

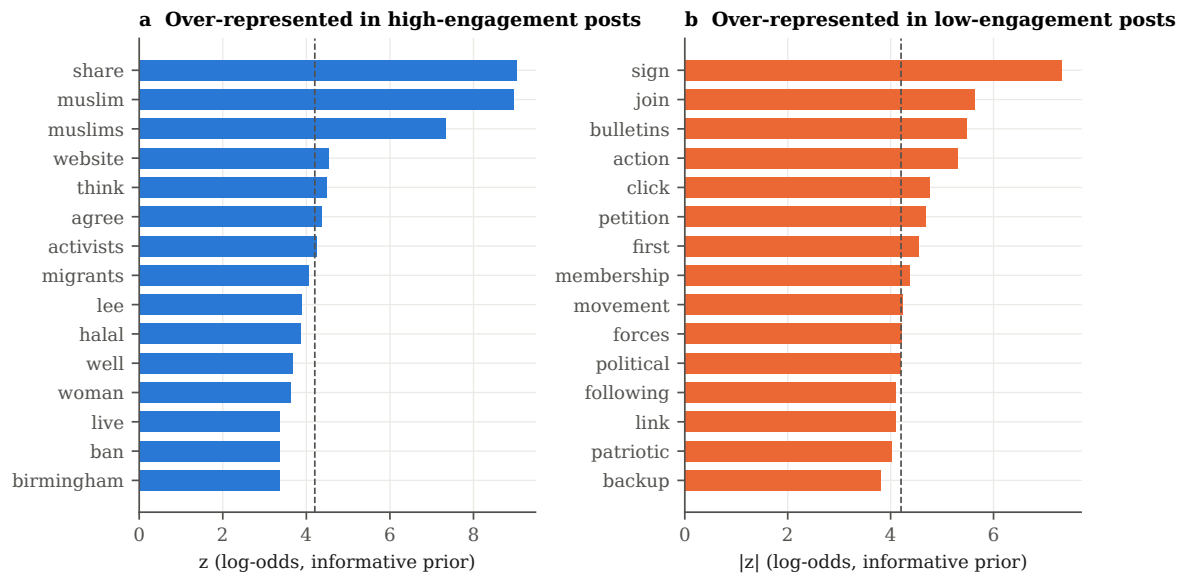


Figure 7. Words associated with engagement. Weighted log-odds z-scores with an informative prior, comparing the top and bottom quartile of posts by engagement net of month and format, excluding auto-generated posts ($n = 841$ and 841). Dashed lines mark the 5% family-wise threshold from a permutation test. “Lee” refers to posts about the 2013 murder of Lee Rigby.

5.7 Reactions: a page built on anger

On the average post published after 24 February 2016, 28% of reactions were angry, 6% haha and 3% love, with the rest like, sad or wow (Figure 8). The angry share varies with the content of the post. Posts about Islam attract an angry share of 36% against 20% for other posts (a difference of 16 percentage points), posts about immigration 35% against 26%, and posts about crime 31% against 27% (all $p < .001$ after Benjamini–Hochberg correction). H3 is supported. In these unadjusted comparisons, posts that carry fundraising or activism content or name the leaders have a lower angry share. With month fixed effects and the other codes, the lower angry share holds for posts naming the leaders (−9 points) but not for fundraising, which has a higher angry share (+12 points, on 21 posts). Within months, the angry share of reactions is correlated with the number of shares ($\rho = 0.27$) and comments ($\rho = 0.36$) a post receives. The data do not show who reacted: angry reactions may come from supporters angry at the subject of a post or from opponents angry at the page, and widely shared posts reach both. The monthly angry share varies from 7% to 42% and fell in late 2017 and early 2018 (Figure 8b), a period in which posts about the leaders and their prosecution became common.

5.8 Comparison with the mainstream parties

How does the audience’s interaction profile compare to that of the mainstream parties’ audiences (RQ1)? Table 4 and Figure 9 report differences in the composition of engagement within month-by-format cells, based on 21 to 58 matched cells per comparison.

Britain First’s audience commented less than the audiences of the Conservative (−21 points), Labour (−10 points) and Liberal Democrat (−9 points) pages. Comments make up 16% of Britain First engagement in the Conservative-matched cells and 36% of Conservative engagement. Britain First’s audience commented more than the Green audience (+5 points) and at about the same rate as UKIP’s (−2 points, a bootstrap interval that includes zero). It shared more than the audience of every other

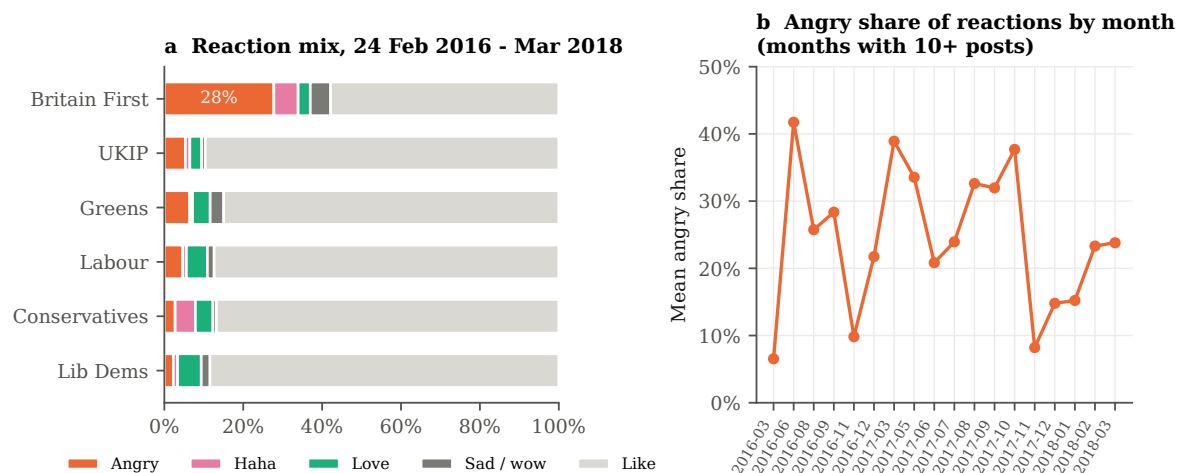
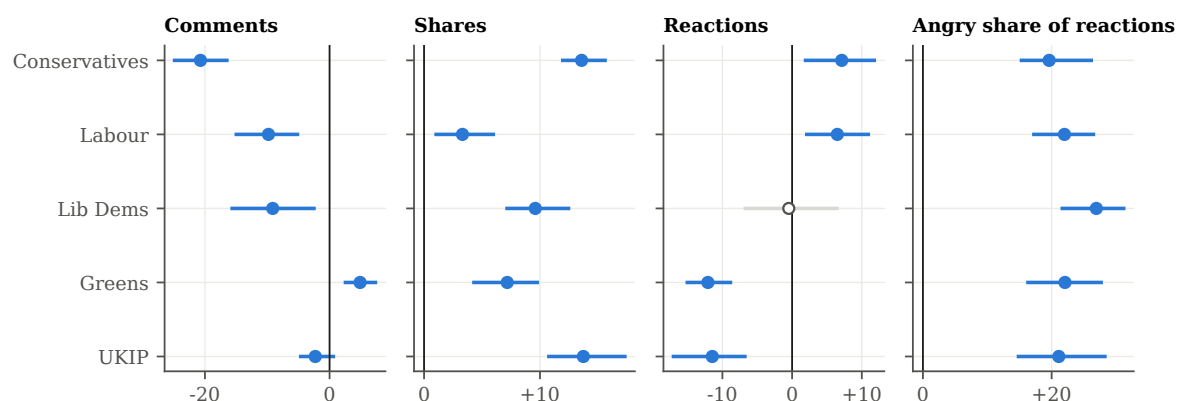


Figure 8. Reactions. Panel a shows the mean composition of reactions by page for posts published from 24 February 2016 to March 2018, unmatched. Panel b shows the monthly mean angry share of Britain First reactions.



Britain First minus comparator page, percentage points of post engagement (95% CI; matched on month and format)

Figure 9. Britain First minus comparator page, in percentage points of post engagement (comments, shares, reactions) or of reactions (angry), matched on month and format, with 95% month-cluster bootstrap intervals. Filled markers pass Holm correction.

page, by 3 to 14 points, with the largest gaps against UKIP and the Conservatives. On reactions as a share of engagement, Britain First is higher than Labour and the Conservatives, lower than UKIP and the Greens, and not different from the Liberal Democrats. Angry reactions made up 20 to 27 points more of its reactions than of any comparator's. Every difference in Table 4 survives Holm correction across the 30 tests I ran (the 20 shown plus haha and love shares), except for reaction share against the Liberal Democrats. Haha share against the Conservatives, which is not shown, also does not survive (a difference of 0.0 points). The permutation test treats posts within a cell as exchangeable and is likely to be anti-conservative: for comments against UKIP it gives $p < .01$ although the month-cluster bootstrap interval includes zero, and I rely on the intervals.

Britain First's engagement leaned towards shares and angry reactions, while the mainstream pages, and especially the Conservative page, drew a larger share of comments. This comparison is of composition and not of scale. I make no claim about relative audience size, which the retrieval constraint prevents me from measuring.

Table 4. Britain First compared with mainstream party pages, matched on month and format (photo, video). Means are percentages of post engagement (comments, shares, reactions) or of reactions (angry), weighted by cell. Differences are in percentage points with 95% month-cluster bootstrap intervals. *p* values come from within-cell permutation tests with Holm correction across 30 tests, shown as <.01 at the resolution of 4,000 permutations (angry share uses posts after 24 February 2016 only).

Page	Measure	Britain First	Comparator	Difference [95% CI]	<i>p</i> (Holm)
Conservatives	Comments	15.7	36.5	-20.7 [-24.9, -16.5]	<.01
	Shares	25.1	11.5	+13.6 [+12.0, +15.6]	<.01
	Reactions	59.1	52.0	+7.1 [+1.9, +11.8]	<.01
	Angry	23.0	3.4	+19.6 [+15.3, +26.1]	<.01
Labour	Comments	15.6	25.4	-9.8 [-15.0, -5.2]	<.01
	Shares	27.0	23.7	+3.3 [+1.0, +6.0]	<.01
	Reactions	57.4	50.9	+6.5 [+2.1, +10.9]	<.01
	Angry	27.7	5.8	+22.0 [+17.2, +26.5]	<.01
Lib Dems	Comments	16.7	25.8	-9.1 [-15.6, -2.5]	<.01
	Shares	27.2	17.6	+9.6 [+7.2, +12.5]	<.01
	Reactions	56.1	56.6	-0.5 [-6.7, +6.4]	1
	Angry	28.5	1.5	+26.9 [+21.7, +31.2]	<.01
Greens	Comments	15.9	11.0	+4.9 [+2.6, +7.4]	<.01
	Shares	27.7	20.6	+7.2 [+4.3, +9.8]	<.01
	Reactions	56.3	68.4	-12.1 [-15.0, -8.8]	<.01
	Angry	25.2	3.1	+22.1 [+16.3, +27.7]	<.01
UKIP	Comments	15.0	17.3	-2.3 [-4.6, +0.6]	<.01
	Shares	27.3	13.6	+13.7 [+10.8, +17.3]	<.01
	Reactions	57.6	69.1	-11.5 [-17.0, -6.8]	<.01
	Angry	24.2	3.1	+21.1 [+14.8, +28.3]	<.01

5.9 Events, timing and frequency

I find no evidence for H7 (Figure 10a; Appendix Table 7). Engagement in the week from the murder of Jo Cox, the EU referendum, the Westminster attack, the Manchester Arena bombing, the London Bridge attack, the Parsons Green bombing and the Trump re-posts did not differ significantly from the surrounding 60 days (all permutation $p > .15$). These tests have low power, since the archive holds few posts in some windows, and the two windows with the most posts show no increase: engagement fell in the weeks of the murder of Jo Cox (-9%) and of the Westminster attack (-27%). The Jo Cox and referendum windows overlap. I read this as no evidence of an event effect, not as evidence that events had none.

I also find no support for H8. The Spearman correlation between the number of posts on a day and the mean engagement of that day's posts, net of month, is -0.04 ($p = .22$, $n = 757$ days), and a regression of mean log engagement on log posts per day with month fixed effects gives a coefficient of -0.050 (95% CI -0.131 to 0.030). Time of day does not differ significantly in a Kruskal-Wallis test net of month ($H = 27.6$, $p = .23$).

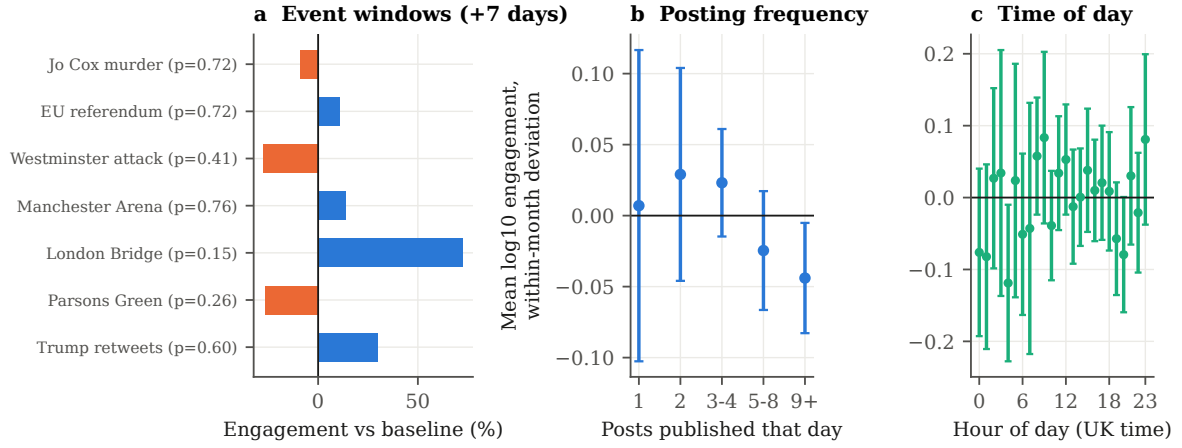


Figure 10. Timing. Panel a shows engagement in the seven days from each event against the surrounding 60 days. Panel b shows within-month mean engagement by number of posts published that day. Panel c shows within-month mean engagement by hour of day (UK time). Bars and whiskers show 95% confidence intervals.

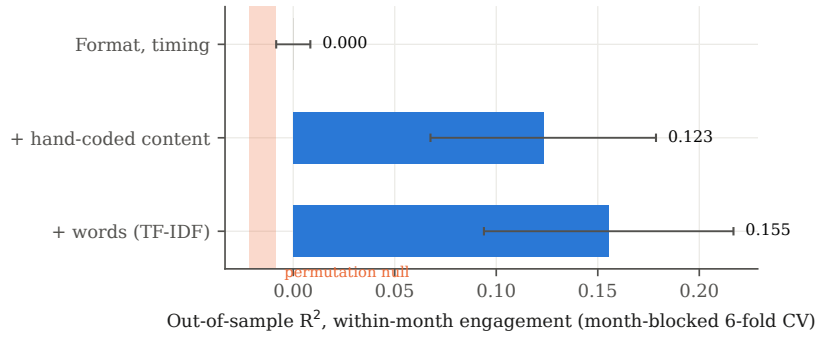


Figure 11. Out-of-sample prediction of within-month engagement (month-blocked six-fold cross-validation). The shaded band shows the range of a permutation null.

5.10 How much of engagement can be predicted from text?

The content codes and the text predict a meaningful share of the variation between posts within the same month. In month-blocked six-fold cross-validation, format and timing predict none of it (out-of-sample $R^2 = 0.000$), hand-coded content features predict 0.123, and adding the words of the post raises it to 0.155 (Figure 11). A permutation null that shuffles engagement within months yields an R^2 of -0.014 on average and, in 12 permutations, never above -0.008 . The text therefore carries information about engagement that generalises across months. They are also modest in absolute terms: most of the variation in what goes viral remains unexplained by the content of the message, which is what the extreme tail in Figure 2 would lead us to expect.

6 Robustness

Table 5 reports the key coefficients across seven specifications: the main model; dropping November and December 2013, the high-volume opening months of the archive; dropping auto-generated posts; winsorising engagement at the 1st and 99th percentiles; restricting to posts after the introduction of reactions; dropping the three months with the most posts; and a model without month fixed effects. The

Table 5. Robustness of key associations. Percentage differences in engagement from the main specification (column 1) and six alternatives. Columns 2–6 use month fixed effects with clustered standard errors; column 7 omits month fixed effects and uses HC3 standard errors. * $p < .05$, ** $p < .01$, *** $p < .001$.

	Main	Excl. Nov–Dec 2013	Excl. auto posts	Winsorised	Post Feb 2016	Excl. 3 largest months	No month FE
Islam	+56***	+62***	+56***	+56***	+75***	+64***	+63***
Immigration	+39***	+43***	+38***	+39***	+43***	+45***	+50***
Crime	+20**	+20**	+20**	+18**	+12	+16*	+33***
Negative sentiment	+14*	+16***	+14*	+14*	+12*	+10	+5
Question	+137***	+150***	+137***	+121***	+200***	+145***	+115***
Share request	+116***	+81***	+113***	+111***	+204**	+90***	-5
Fundraising	-58***	-48**	-58***	-57***	-15	-56***	-34*
URL	-51***	-66***	-51***	-50**	-78***	-54**	-64***
Link format	-53**	-38	-57***	-53**	-26	-48*	-60***
Video	+1	-4	+4	+1	-2	+4	+3
Own leaders	+6	+8	+7	+6	0	+2	+26**
<i>n</i>	3505	2953	3361	3505	1783	2673	3505

hostile-theme associations (Islam, immigration, crime), the question and share-request associations and the penalties for links, URLs and fundraising keep their sign across all of them, although the crime association and the fundraising and link penalties are smaller and not significant in the post-February 2016 subsample. Two results require caution. The share-request association is not recovered without month fixed effects (-5%), where it is confounded with the page’s early growth. A Poisson model on raw counts gives disproportionate weight to a few viral posts and is imprecise (Appendix Table 8). The question association holds (incidence rate ratio 3.97, $p = .007$), as does crime (2.02, $p = .048$), and the Islam estimate is positive (1.44, $p = .09$), but the immigration estimate is not. The share-request estimate reverses sign, because the two posts with over 3.8 million interactions, both from the page’s first seven months, dominate raw counts. Quantile regression shows that the Islam, immigration and question associations hold at both the median and the 90th percentile, and that the crime association holds at the median only.

7 Discussion

The page’s audience engaged most with content about out-groups and with direct invitations to react, and least with the organisational work. I draw three implications.

First, on the question of what a far-right page learns from its audience, the associations point in one direction. Nearly every association in the data favours more content about Islam, immigration and crime and more interactive prompts, and disfavours the content that asks for membership and money. Among the retrieved posts, Islam-related posts rose from 15% in late 2013 to 64% in 2016 before falling to 31% in early 2018; because the archive holds different months in each year, these yearly shares are not a clean trend. I cannot show that the page’s editors responded to engagement signals, because the archive contains no information about its internal decisions. The match between what was rewarded and what was posted is consistent with that mechanism and with other explanations, including a straightforward ideological commitment.

Second, the finding that engagement and mobilisation diverge matters for how we read engagement data from extremist and other politically active pages. A page that has millions of interactions has not necessarily built a movement. In this archive, the content its audience engaged with least is the

content that would have built an organisation: petitions, memberships, donations and links to its own site. The result is consistent with the claim that attention on a platform and political capacity are different resources, and that a strategy of maximising the first can crowd out the second. It also offers one reading of why a movement whose posts drew millions of interactions won no electoral foothold, although the archive cannot test that claim.

Third, the absence of a detectable leadership effect and the high share of shares in engagement suggest that the page worked more as a distributor of issue content than as a leader-centred community. For platform governance, the implication is that action against specific leaders and pages, as Facebook took in March 2018, removes a distribution node but not the themes and prompts that, in these data, the audience rewarded and that any actor can reproduce.

The comparison with mainstream parties shows the same difference. Conservative and Labour posts drew comment shares 1.6 to 2.3 times those of Britain First's, while Britain First's posts drew more shares and more angry reactions. These are descriptive differences within matched cells, and they describe what the audiences did, not their motives.

8 Limitations

The archive is a retrieval-limited subset, not a census, and it may over-represent high-engagement posts. I handle this with month fixed effects, with within-month inference and with a truncation sensitivity analysis, but I cannot recover the posts that were not retrieved. Some months have no posts at all, and monthly coverage is uneven, which limits my ability to describe trends and to detect event effects. The coding is dictionary-based and applies to message text only. It misses images, sarcasm and implicit references, so some categories (leader usage in particular) are underestimated, and measurement error in the content codes attenuates coefficients toward zero. The associations are observational, and I cannot rule out omitted variables that vary within a month, such as the content of images. Facebook's ranking algorithm changed over the period, so engagement reflects both audience behaviour and algorithmic distribution, and the data cannot separate them. I could not observe reach or impressions, only interactions. Finally, the comparison pages are subject to the same retrieval limits, and I compare composition, not scale. Throughout, I use "associated with" where the design supports only association.

9 Conclusion

An archive that cannot be rebuilt shows how the audience of a prominent far-right Facebook page behaved over four years. Engagement was extremely concentrated in a few viral posts. It was higher for posts about Islam, immigration and crime, and for questions and requests to share, and lower for posts that asked supporters to join, sign or give. The audience shared and reacted angrily more than mainstream party audiences did. Naming the leaders made no measurable difference. For researchers, engagement on this page tracked attention to out-group content and prompts and did not track the content that asked for organisational commitment. For platforms and regulators, a ranking system that rewards engagement would have favoured the out-group content that this page's audience engaged with most. The wider body of data from the period, covering the mainstream parties and their leaders, offers the chance to test whether the reward structure found here is specific to the far right or a general property of the platform. That question is the natural next step for this body of data.

Data and code availability. The coded post-level dataset (post identifiers, dates, formats, codes and engagement counts, without message text) and the analysis code are available from the author on request.

References

- Bene, M. (2017). Go viral on the Facebook! interactions between candidates and followers on Facebook during the Hungarian general election campaign of 2014. *Information, Communication & Society*, 20(4):513–529.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1):289–300.
- Berger, J. and Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2):192–205.
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., and Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28):7313–7318.
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.
- Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, 114(3):494–509.
- Engesser, S., Ernst, N., Esser, F., and Büchel, F. (2017). Populism and social media: How politicians spread a fragmented ideology. *Information, Communication & Society*, 20(8):1109–1126.
- Ernst, N., Engesser, S., Büchel, F., Blassnig, S., and Esser, F. (2017). Extreme parties and populism: An analysis of Facebook and Twitter across six countries. *Information, Communication & Society*, 20(9):1347–1364.
- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4):665–668.
- Gerbaudo, P. (2018). Social media and populism: An elective affinity? *Media, Culture & Society*, 40(5):745–753.
- Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70.
- Hutto, C. J. and Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*.
- ITV News (2017). Theresa May tells Donald Trump: “retweeting Britain First was the wrong thing to do”. ITV News, 29 November 2017.

- ITV News (2018). Facebook bans Britain First pages. ITV News, 14 March 2018.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.
- Monroe, B. L., Colaresi, M. P., and Quinn, K. M. (2008). Fightin’ words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403.
- Mudde, C. (2007). *Populist Radical Right Parties in Europe*. Cambridge University Press, Cambridge.
- Rathje, S., Van Bavel, J. J., and van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences*, 118(26):e2024292118.
- Rieder, B. (2013). Studying Facebook via data extraction: The Netvizz application. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 346–355.
- Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

A Appendix

A.1 Truncation sensitivity

Table 6. Truncation sensitivity. Percentage differences in engagement from the main specification after dropping the lowest-engagement 0%, 10%, 25% and 40% of posts within each year. Shrinking estimates as the lower tail is removed indicate that selection of posts from below attenuates the associations; the estimates for the full archive are then conservative. * $p < .05$, ** $p < .01$, *** $p < .001$.

	Full archive	Drop lowest 10%	Drop lowest 25%	Drop lowest 40%
Islam	+56***	+37***	+26***	+13*
Immigration	+39***	+27***	+18*	+9
Crime	+20**	+19**	+13	+14*
Negative sentiment	+14*	+11*	+12**	+7
Question	+137***	+99***	+103***	+77***
Share request	+116***	+84***	+68***	+53**
Fundraising	-58***	-22	+3	+23
URL	-51***	-42**	-26	-16
Link format	-53**	-35**	-30	-16
Video	+1	-12	-13	-9
Own leaders	+6	+9	+11	+3
<i>n</i>	3,505	3,155	2,631	2,104

A.2 Event windows

Table 7. Event windows. Engagement in the seven days from each event (day 0 to day 7) against posts in the 60 days before and after the window, net of post format. p values come from a day-level permutation test with 5,000 permutations.

Event	Date	Posts in window	Posts in baseline	Difference (%)	p
Jo Cox murder	2016-06-16	73	180	-9	0.72
EU referendum	2016-06-23	41	214	+11	0.72
Westminster attack	2017-03-22	60	151	-27	0.41
Manchester Arena	2017-05-22	41	115	+14	0.76
London Bridge	2017-06-03	13	118	+73	0.15
Parsons Green	2017-09-15	28	237	-26	0.26
Trump retweets	2017-11-29	5	265	+30	0.60

A.3 Poisson pseudo-maximum-likelihood estimates

Table 8. Poisson pseudo-maximum-likelihood estimates on raw engagement counts with month fixed effects and month-clustered standard errors (incidence rate ratios). Estimates are dominated by a few viral posts.

	IRR	95% CI	<i>p</i>
Islam	1.44	[0.94, 2.20]	.090
Immigration	0.89	[0.66, 1.21]	.476
Crime	2.02	[1.01, 4.04]	.048
Negative sentiment	1.04	[0.63, 1.72]	.878
Question	3.97	[1.46, 10.81]	.007
Share request	0.23	[0.05, 1.07]	.061
Imperative verb	2.11	[1.10, 4.07]	.025
Own leaders	1.30	[0.65, 2.60]	.451
Video	1.14	[0.58, 2.23]	.707
Link format	0.49	[0.19, 1.32]	.159

Provenance and AI use

Note added in 2026 to version 2 of this preprint. The rest of the paper is unchanged from version 1.

The data were collected with Netvizz during my doctoral research. This paper builds on my doctoral thesis (Hotham 2021). The analysis and writing were completed in 2026.

The analysis code and drafts of the text were prepared with the assistance of an AI system (Claude, by Anthropic), which was also used to review drafts. I checked the analysis and the text and take responsibility for both.

References for this note

Hotham, T. (2021). *How Do Political Parties Use Facebook and What Does It Offer to Their Campaigns*. PhD thesis, University of Bath.