CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Misremembering Brexit: partisan bias and individual predictors of false memories for fake news stories among Brexit voters
Authors	Greene, Ciara M.;Nash, Robert A.;Murphy, Gillian
Publication date	2021-05-10
Original Citation	Greene, C. M., Nash, R. A. and Murphy, G. (2021) 'Misremembering Brexit: partisan bias and individual predictors of false memories for fake news stories among Brexit voters', Memory, 29(5), pp. 587-604. doi: 10.1080/09658211.2021.1923754
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1080/09658211.2021.1923754
Rights	© 2021, Informa UK Limited, trading as Taylor & Francis Group. This is an Accepted Manuscript of an item published by Taylor & Francis in Memory on 10 May 2021, available online: https://doi.org/10.1080/09658211.2021.1923754 - https://creativecommons.org/licenses/by-nc/4.0/
Download date	2026-09-10 04:28:10
Item downloaded from	https://hdl.handle.net/10468/14244

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Misremembering Brexit: Partisan bias and individual predictors of false memories for fake news stories among Brexit voters

Ciara M. Greene[1*], Robert A. Nash[2] and Gillian Murphy[3]

[1] School of Psychology, University College Dublin, Ireland

[2] School of Psychology, Aston University, Birmingham, UK

[3] School of Applied Psychology, University College Cork, Ireland

*Corresponding author. Email ciara.greene@ucd.ie

ABSTRACT

Exposure to fake news stories can result in false memories for the events portrayed, and this effect can be enhanced if the stories conform to the reader's ideological position. We exposed 1299 UK residents to fabricated news stories about Brexit. 44% of participants reported a false memory for at least one fabricated story, with a higher rate of false memories for stories that reflected poorly on the opposing side. This effect of ideological congruency was somewhat greater among participants who were exposed to a threat to their social identity as a Leave or Remain supporter; however, this moderating effect was only statistically significant in exploratory analyses using a more conservative definition of false memory. Participants with higher cognitive ability and analytical reasoning scores were less susceptible to false memories. Individuals with better knowledge about Brexit showed better discrimination between true and false stories, while self-reported engagement with the Brexit debate was associated with an increased tendency to 'remember' any story, regardless of its truth. These results implicate a combination of social and individual factors in the development of false memories from fake news, and suggest that exposure to social identity threats may enhance the polarising effects of fake news.

Keywords: False memory; fake news; motivated reasoning; analytical reasoning; engagement

Word count: 10,895

INTRODUCTION

Although news outlets have always circulated a certain amount of biased or misleading information, the rise of the internet in the late 20th and early 21st century has sharply increased the quantity and accessibility of misleading and inaccurate information (Lazer et al., 2018). Recent research has shown that when people see fabricated news stories or photographs of public events that never took place, they can come to believe in and even remember those fictional events (Nash, 2018; O'Connell & Greene, 2017; Sacchi et al., 2007; Strange et al., 2011), especially if the content of the fabricated material is consistent with their political or ideological views (Frenda et al., 2013; Murphy et al., 2019). Implantation of false memories can affect subsequent behaviour (see Laney & Loftus, 2017 for a review), and recent evidence suggests that the formation of a false memory for a fake news story can influence related behavioural intentions, over and above mere exposure to the fabricated story (Greene & Murphy, 2020). Thus, investigating factors that affect the development of false memories for fake news is an important scientific endeavour. In this paper we present a wide-ranging investigation of individual and social factors that influence the formation of false memories for partisan fake news. As a context for achieving this goal, we capitalise upon the polarised political climate in the United Kingdom following the country's decision in 2016 to leave the European Union.

False memories

One influential approach to understanding false memory generation is the source monitoring framework (Johnson et al., 1993; Mitchell & Johnson, 2000). According to this approach, people distinguish true from false memories by evaluating their mental experiences on the basis of two cognitive systems, operating in parallel. The first relies on heuristic judgements about the perceptual, temporal, and semantic features of the putative memory to rapidly determine whether they match the typical characteristics of a true memory. For example, true memories are typically more vivid and familiar, and contain more spatial and temporal details than imagined or otherwise non-veridical

experiences. The second system relies on a more deliberative, systematic process whereby the contents of the putative memory are carefully assessed for plausibility, or compared against prior knowledge and other (dis)confirmatory evidence. Both heuristic and systematic processes can lead people to incorrectly attribute internally-generated mental experiences to real memories of past events. For example, this may occur when an eyewitness incorrectly decides that a mental image of a crime arises from their memory of witnessing the event rather than from the prosecutor's leading question (Lindsay, 1994). Following this initial judgement, the individual may go on to confabulate details of the event and construct a false memory (Hyman & Kleinknecht, 1999; Mazzoni et al., 2001; Mitchell & Johnson, 2000; Strange et al., 2005).

The source monitoring framework can be readily applied to the cognitive processing of fake news, where, for example, a news consumer encountering a fabricated story may quickly determine that the story *feels* true, or may evaluate the story's plausibility in light of their own experiences, knowledge and preconceptions, and determine that the story is *likely* to be true (Mazzoni & Kirsch, 2002). A body of research supports the idea that perceived plausibility is an important precursor to believing in an event's occurrence, although what individuals consider plausible can itself be influenced by their prior knowledge and experiences (Mazzoni et al., 2001; Scoboria et al., 2007; Scoboria et al., 2004). Judgments such as these may lead the individual to draw on their existing knowledge and memories of related news stories to construct a detailed memory of the fictional event (O'Connell & Greene, 2017).

Motivated reasoning

Not all fake news is created equal, and in line with the source monitoring framework's principles, research demonstrates that we are particularly susceptible to forming false memories for stories that align with our pre-existing beliefs or political orientations. For example, Frenda et al. (2013) reported that American liberals were more likely to recall a fabricated scandal about President Bush, while

conservatives were more likely to recall a fabricated scandal about President Obama. Similarly, within the context of a national referendum on abortion, Murphy et al. (2019) reported that participants tended to form false memories for political scandals that aligned with their stated position. Moreover, many participants reported detailed false memories of these fictitious political events, and expressed strong emotional responses to them.

Decades of research inspired by social identity theory (Tajfel & Turner, 1979) suggests that, once categorised into a group, individuals assimilate membership of that group into their identity, resulting in high levels of in-group identification and out-group derogation (see Hornsey, 2008 for a review). A direct corollary of this principle is that strong identification with an ideological position or political party does not require a sophisticated understanding of the ideology itself; instead, political ideology derives much of its behavioural influence from social identification with the members of a group espousing that ideology (Devine, 2015). These social identities can have dramatic effects on downstream reasoning and behaviour; for example, Cohen (2003) found that attitudes towards a social policy were overwhelmingly determined by which political party was said to have proposed the policy, irrespective of its objective content or the participant's own ideological beliefs.

Psychological scientists have repeatedly suggested that this partisanship leads to motivated reasoning, which advances the interests of the in-group by attending to identity-congruent information while disregarding incongruent details (Bolsen et al., 2014; Kahan, 2015; Schaller, 1992). As a result, belief in and response to partisan fake news is frequently distorted by social identity goals, to the detriment of accuracy goals (Kahan, 2017; Van Bavel & Pereira, 2018). The effects of this distortion can be seen in the formation of attitudes on polarising topics such as climate change (Druckman & McGrath, 2019; Hart & Nisbet, 2012) or immigration (Druckman et al., 2013), where identification with a political party or ideology is the main driver of attitude formation. Since false memories often follow false beliefs (Hyman & Kleinknecht, 1999; Mazzoni et al., 2001; Scoboria et al.,

2004), motivated reasoning may explain why people are more likely to form false memories for fake news stories that support their political position. In the interests of simplicity, we hereafter refer to fabricated news stories that align with an individual's social and political identity as "ideologically congruent" fake news.

Presenting a threat to an individual's social identity can magnify motivated reasoning effects (Branscombe et al., 1999). There are many different types of social identity threat, including threats to the existence of the group and threats to the individual's group membership; in the present study we focus on threats to the *value* of the group or social identity, where the group's competence or morality is called into question. This form of threat typically results in increased in-group identification and differentiation from the outgroup (Branscombe et al., 1999; Grant & Brown, 1995; Voci, 2006). Individuals who identify strongly with the group identity tend to derogate the outgroup even more strongly after their own group's value is threatened (Branscombe & Wann, 1994). If the formation of false memories following exposure to ideologically congruent fake news can be attributed to motivated reasoning on the basis of group identity, then we might expect these effects to be enhanced if participants are first exposed to a threat to that identity. The presence of a threat would be expected to result in increased false memories for fabricated stories that derogate the out-group. In the present study, we investigate whether an identity threat directed at Leave or Remain voters in the 2016 Brexit referendum affects formation of false memories for ideologically congruent fake news.

Cognitive ability and reasoning style

In addition to the social and group-level factors outlined above, individual differences in a range of cognitive variables have been implicated in susceptibility to fake news. Murphy et al. (2019) reported that false memories for ideologically congruent fake news relating to the Irish abortion referendum were observed most frequently amongst individuals with lower scores on the Wordsum, a measure of cognitive ability that is highly correlated with full-scale IQ (Huang & Hauser, 1996; Thorndike & Gallup,

1944). This finding suggests that false memory formation arising from exposure to fake news may be tempered by a rational or intellectual approach to information processing. Converging evidence from eyewitness memory studies suggests that people with higher levels of cognitive ability and working memory capacity may be less susceptible to misinformation, as these individuals are better equipped to compare new memories to old, and discard inconsistent information (Greene et al., 2020; Jaschinski & Wentura, 2002; Lee, 2004; Zhu et al., 2010).

The propensity to form false memories from fake news may also be influenced by an individual's reasoning style. Pennycook and Rand (2019) reported that belief in fake news was best explained by participants' performance on the Cognitive Reflection Test (CRT; Frederick, 2005). They argued that participants who perform poorly on this test tend to reason intuitively rather than analytically, and thus may be more likely to take fake news stories at face value, rather than critically evaluating their source and plausibility (see also Bronstein et al., 2019; Pennycook & Rand, 2018). Hence, individuals who perform poorly on the CRT may be more prone to forming false memories when exposed to fake news. CRT performance is moderately correlated with full-scale IQ and with the Wordsum (Murphy et al., 2019; Toplak et al., 2014), suggesting that although these measures share some variance, they are tapping different underlying cognitive constructs. In the present study, we investigated the contribution of both general cognitive ability (assessed using the Wordsum) and analytical reasoning (assessed using the CRT) to false memory formation.

Subject knowledge, interest and engagement

A combination of social (ideological congruency, identity threat) and individual variables (cognitive ability, reasoning style) have thus far been identified as potential contributors to the development of false memories for fake news. However, despite similarities in their social or ideological identities, individuals also vary in their past exposure to and engagement with political debate. A news consumer's response to a particular story is likely to be informed by their past experience with similar

content; indeed, the nature and quality of an individual's past exposure to and engagement with the topic under discussion may be an important determinant of whether an individual falls for fake news about that topic (O'Connell & Greene, 2017).

Experts are sometimes more prone than non-experts to developing false memories in relation to their area of expertise (Baird, 2003; Castel et al., 2007; Mehta et al., 2011). This finding has been attributed to experts' more developed schemata for their subject area (Baird, 2003); however even having a strong interest in a topic has been shown to double the rate of false memories, relative to a low-interest topic (O'Connell & Greene, 2017). From a source monitoring perspective, this effect may occur because individuals with considerable experience in a given topic have multiple overlapping memory traces related to that topic; hence, a novel stimulus relating to that topic is more likely to activate a similar memory trace, triggering a sense of familiarity followed by the construction of a false memory (Mitchell & Johnson, 2000).

A competing explanation is provided by Mehta et al. (2011), who reported that people who self-identify as experts in a particular topic feel accountable for their ability to recall information pertinent to that topic, and that this accountability mediates the experts' level of false recall. False memory effects among experts may therefore be attributed to a tendency towards overclaiming, arising from a fear of appearing ignorant. The effects of interest in, engagement with and knowledge of a particular topic are hard to disentangle; an individual who is interested in a given topic may engage with more online and offline content related to that topic; and depending on the sources they engage with, they may become more knowledgeable on the topic. This may result in both a greater number of overlapping memory traces, and an increased desire to present themselves as knowledgeable on the subject. Moreover, expertise and knowledge of a topic are often self-reported, rather than objectively assessed, and people who overclaim their knowledge of a topic are more likely to accept fake news headlines as accurate (Pennycook & Rand, 2018), further complicating the task of distinguishing the

effects of interest, engagement and knowledge. In the present study, we evaluated knowledge about Brexit using a multiple-choice test, and also assessed self-reported interest and engagement with the Brexit debate.

To summarise, while extensive work has been conducted on responses to politically partisan misinformation, the effects of a number of social and individual factors have yet to be elucidated. These include the influence of social identity threats and individual differences in cognitive style and ability, as well as subject-specific knowledge and interest. Here, we describe a comprehensive investigation of a range of these factors in the context of the Brexit referendum. In addition to analysing the occurrence of false memories as a function of these factors, we report exploratory analyses using signal detection measures of response bias and sensitivity. The basic premise of signal detection theory (Green & Swets, 1966) is that an observer searching for a signal in a noisy field will report a signal as present if their internal representation of the signal rises above a given threshold, and absent if it does not reach the threshold. Thus, the observer's response may be categorised based on a combination of the presence or absence of the signal and the observer's *report* of the presence or absence of the signal. In the context of the present study, if the 'signal' is a veridical memory of a real event, then reported memories of true stories may be considered hits while reported memories of fabricated events may be considered false alarms. These measures might be used to distinguish between participants' tendency to overclaim (i.e. their response bias) and their ability to discriminate true from false stories (i.e. sensitivity).

Brexit and the present study

The 2016 Brexit referendum, and subsequent debate in the UK, provide an ideal context for investigating the questions outlined above. On the 23rd of June 2016, the United Kingdom voted to leave the European Union. The process of negotiating a withdrawal agreement was protracted, and the UK finally left the EU on the 31st of January 2020, marking the start of a lengthy transition period

involving further political debate and negotiations. Turnout on Brexit polling day was high (72.2%) and the vote was very close, with 52% of the electorate voting Leave, and 48% voting to remain in the EU (BBC News, 2016). The issue remains contentious, and the group identities of "Leaver" and "Remainer" have taken on persistent political and social meanings that continue to influence political decision-making (Axe-Browne & Hansen, 2020; Hobolt et al., 2020). The Brexit debate therefore affords a binary social identity in the UK, and news reports and analysis in the British media frequently contain material that is critical of either Leavers or Remainers (Hinde, 2017); consider for example, the numerous unflattering references in tabloid newspapers to "Remoaners" – Remain supporters who were unhappy with the outcome of the referendum. Thus, members of the public are frequently subject to myriad threats to their social identity.

In the present study, we presented British participants with a social identity threat in the form of a fabricated television news report that included derogatory comments about either the Leave or Remain side in the Brexit debate. We then assessed participants' memory for a number of true and false news stories that related to the events surrounding Brexit. Finally, we collected measures of analytical reasoning, cognitive ability, and general knowledge in relation to Brexit, as well as self-report measures assessing degree of interest and engagement in the Brexit debate. Our preregistered hypotheses were as follows:

1. That participants would form more false memories for Brexit-related fake news stories that were ideologically congruent than for stories that were ideologically incongruent
2. That the effect of ideological congruency would be enhanced when a threat was presented to the individuals' group identity as a Leave or Remain supporter, compared with a threat to the opposing group's identity
3. That participants with higher scores on measures of cognitive ability and analytical reasoning would form fewer false memories for Brexit-related fake news stories

4. That participants with higher levels of knowledge, interest and engagement about Brexit would form more false memories for Brexit-related fake news stories.

METHOD

Preregistration

We separately preregistered the two arms of the study, concerning (1) effects of ideological congruency and identity threat, and (2) individual differences in false memory formation, at https://aspredicted.org/52fn9.pdf and https://aspredicted.org/ub3fm.pdf, respectively. We preregistered a sample size of 1300 participants, based on the results of power analysis focused on the two primary analyses in this paper, namely:

1. Two-way between-subjects ANOVA: 1302 participants required to detect small effects ($f = 0.1$) with 95% power and an alpha level of 0.05
2. Multiple linear regression: 995 participants required to detect small effects ($f^2 = 0.02$) with 95% power in a model with five predictors and an alpha level of 0.05.

Power analyses were conducted in G*Power version 3.1 (Faul et al., 2007).

The study protocol was approved by the Human Research Ethics Committee (Humanities) of University College Dublin.

Participants and Design

Participants were recruited via the online crowdsourcing platform Prolific (https://www.prolific.co/) and via social media and personal networks. Prolific participants were paid £1.25 on completion of the study. Participants were required to be over 18 and a UK national. The study was completed by 1533 individuals; participants who failed an audio check (N = 25) or two or more attention checks (N = 209) were excluded in accordance with our preregistered criteria, leaving 1299 participants (454 male, 839 female, 6 other; mean age 35.88 years, SD = 12.43, range = 18-76). The final sample size is in line with our preregistered target of 1300. Each participant was randomly assigned to one of the cells in a 2 (ideological congruency: congruent vs. incongruent fabricated stories) x 2 (threat condition: own group vs. other group) between-subjects design.

Materials

All study materials can be found online at https://osf.io/rn4ae/.

Videos

We created two versions of a video purporting to be an Australian news report about Brexit. Both videos were approximately 2.5 minutes long, and were mostly identical. The identical segments consisted of stock footage of UK landmarks (Big Ben, Trafalgar Square), supermarkets, and street scenes (including Brexit protests, featuring both Leave and Remain supporters), mixed in with footage from a House of Commons debate. A voiceover gave an update on the Brexit negotiations ("No progress on Brexit today, as British MPs continue to debate the Prime Minister's plans for exiting the European Union…") and mentioned concerns about shortages of goods in the event of a "No-Deal" Brexit. The voiceovers for the two versions then diverged, as they described the results of a study conducted by the fictional think tank "Inside Britain". Our aim was to create content that would represent a legitimate and plausible threat to the identity of an avowed Leave or Remain supporter. Thus, rather than presenting identical content about the two groups, we attempted to exploit common criticisms levelled at Leavers and Remainers. In developing these materials, we piloted multiple versions of the video scripts with approximately 160 British Prolific participants (who did not participate in the main study) to identify wording that was considered equally insulting to both sides. The final 'Leave threat' version suggested that Leave voters were less intelligent and less educated than Remain voters, whereas the 'Remain threat' version suggested that Remain voters were elitist and undemocratic for refusing to accept the will of the people. Both videos also included an inflammatory quote supposedly from the lead researcher of the fabricated study, which was read aloud by the reporter and presented onscreen in written form against a background of overlaid EU and UK flags. These quotes can be seen in Figure 1; the full videos and scripts can be found at https://osf.io/rn4ae/.

[Insert Figure 1 here]

Fabricated stories

Two fabricated stories were created, one describing election tampering during the 2016 Brexit referendum, and one describing a Wikileaks report of illegal campaign donations. Identical versions of each story were created which implicated either the Leave or Remain campaign in wrongdoing. The election tampering story read, "The [Vote Leave/Stronger in Europe] campaign was accused of election tampering after 1100 ['Remain'/'Leave'] ballots were found in a skip outside a polling station in Bermondsey one week after the 2016 Brexit referendum". The Wikileaks story read, "Confidential documents released on the WikiLeaks site in October 2017 revealed that nearly 30% of campaign donations to the official [Leave/Remain] campaign came from sources outside the UK, despite strict rules imposed by the Electoral Commission to avoid foreign interference in elections". We conducted in-depth online research to ascertain that events very similar to those depicted in the fabricated stories had not in fact been reported in the media. The stories were pilot-tested with a further 36 participants who did not take part in the main study, to ensure plausibility and to establish that the fictitious events (if true) would reflect poorly on the targeted group. Each story was accompanied by an illustrative but non-probative photograph (for the election tampering story, a photograph of election officials counting votes at a polling station; for the Wikileaks story, a photograph of Wikileaks founder Julian Assange superimposed on the Wikileaks logo). All materials can be viewed online at https://osf.io/rn4ae/.

True stories

Four true stories relating to Brexit were shown to all participants. Matching the format of the false stories, each true story consisted of a short description accompanied by an illustrative photograph. The four true stories were:

1. In September 2019, the Labour leader Jeremy Corbyn was branded a 'gigantic chlorinated chicken' in the House of Commons over his refusal to call for a general election, prompting

Conservative party members to tweet photoshopped images of Mr. Corbyn dressed as a chicken.

2. The value of the pound fell to its lowest level since 1985 after the results of the Brexit referendum were announced, with wild fluctuations observed in international currency markets following the surprise outcome of the referendum.
3. Ford announced in June 2019 that it would be closing its factory in Bridgend, South Wales, with the loss of at least 1500 jobs. The announcement came shortly after the car manufacturer warned that Brexit would be 'catastrophic' for the UK car industry.
4. US vice president Mike Pence raised eyebrows while on a visit to Dublin when he urged Ireland and the EU to deal with the UK "in good faith" during Brexit negotiations.

Brexit knowledge test

We designed a 10-item multiple choice test to assess subject knowledge about Brexit (see online materials). The questions were developed based on issues relating to the Brexit vote and its political aftermath that had been widely documented in news coverage between 2016 and 2020. A sample question from this test is "Which regions of the UK voted to **remain** in the EU? A) Scotland and Northern Ireland; b) Wales, Scotland and Northern Ireland; c) Scotland; d) none". Each question had four possible responses; chance performance is thus 25%. Question order and response order within questions were randomised. The multiple-choice test was pilot-tested with 24 Prolific participants who did not participate in the main study, to ensure comprehensibility of the questions and an appropriate level of difficulty. A mean score of 6.29/10 was achieved in this pilot test (range: 2-9).

Cognitive Ability

Cognitive ability was assessed using the Wordsum, a 10-item subtest of the Wechsler Adult Intelligence Scale vocabulary test (Thorndike & Gallup, 1944). The Wordsum has been reported to be highly correlated with full-scale IQ (Huang & Hauser, 1996) and is used extensively in the US General

Social Survey, where the average score is 6/10 (Meisenberg, 2015). Participants are presented with a target word and asked to choose the word closest in meaning to the target from a list of 5 other words. For example, given the target word ANIMOSITY, participants must choose the closest match from the words 'hatred', 'animation', 'disobedience', 'diversity' and 'friendship', or may select 'don't know'.

Analytical Reasoning

The Cognitive Reflection Test (CRT) is composed of verbal reasoning problems in which each item has an intuitive but incorrect answer, and a correct answer that requires analytical reasoning. A sample question is "A bat and ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?" The intuitive answer to this question is 10c; the correct answer is 5c. The original CRT (Frederick, 2005) consisted of three items that rely on numerical reasoning; the seven-item version used in the present study includes four additional non-numerical items (Thomson & Oppenheimer, 2016; Toplak et al., 2014). A non-numerical example asks, "How many cubic feet of dirt are there in a hole that is 3' deep x 3' wide x 3' long?". The correct answer here is "none". CRT scores are strongly associated with SAT scores, delay discounting and need for cognition (Frederick, 2005), as well as a range of other reasoning measures (Hoppe & Kusterer, 2011; Toplak et al., 2011), and its predictive ability is robust to multiple exposures (Bialek & Pennycook, 2018).

Procedure

Data collection began on 30th January 2020, one day before the EU's official exit from the EU, and continued until 21st March 2020. The data were collected via the online survey platform Qualtrics (https://www.qualtrics.com/). Participants were informed that the purpose of the study was to investigate "attitudes towards the 2016 Brexit Referendum and its aftermath, and how these may change over time". After providing consent, participants supplied demographic details (age, gender, education) and answered a series of questions about how they had voted in the 2016 EU referendum,

how they would vote if the referendum were held tomorrow, how important the outcome of the referendum was to them, and where they would place themselves on a spectrum from liberal to conservative. Participants were also asked to indicate how often in the past three months they had (1) consumed traditional media related to Brexit, (2) engaged with social media content related to Brexit, and (3) had discussions about Brexit with family and friends.

Participants were then randomly assigned to view a video that was insulting to either Leave or Remain voters. To ensure that participants could hear the video, participants were first instructed to listen to a brief recording, which consisted of three beeps followed by the numbers "two, four, eight", and to write down the numbers they heard. They were then instructed to watch the video, and were informed that they would be asked questions about it afterwards. The video was followed by three multiple-choice attention check questions which asked about prominent details from the video (see online materials for further details). In accordance with our preregistered exclusion criteria, participants who failed the audio check and/or two or more attention check questions were excluded from further analysis (see Participants section). A series of Likert scale questions was then presented which asked participants to rate how surprising (1 = not at all surprising, 9 = very surprising), biased (1 = completely biased towards Leave, 9 = completely biased towards Remain) and insulting (1 = insulting to Leave voters, 9 = insulting to Remain voters) the video was.

Participants then viewed six news stories about Brexit, including the four true stories and two fabricated stories. All participants saw the same true stories, but participants were randomly assigned to view fabricated scandals about either the Leave or Remain campaign; all participants saw both the Wikileaks story and the election tampering story about the same campaign (i.e., ideological congruency was manipulated between-subjects). The presentation order of the six stories was randomised for each participant. After each story, participants were asked "Do you remember this event?" and could respond either "I remember seeing/hearing about this", "I don't remember

seeing/hearing about this, but I remember it happening", "I don't remember this, but I believe it happened", "I remember this differently" or "I don't remember this". Participants were then asked where they first saw/heard about the event, and could select a specific source (television, newspaper, radio, online news website, social media, word of mouth, or other source) or could indicate either that they did not see or hear about the event, or that they could not remember where they saw or heard about it. If they remembered the event, participants were asked to indicate via an open-ended text box how they felt about it at the time. Finally, they were asked "Thinking about the event now, how do you think the [leak about campaign donations/story about election tampering] reflected on the [**Leave/Remain**] campaign?" and responded on a scale where 1 = reflected very well and 9 = reflected very badly.

After viewing all six stories, participants completed the Brexit subject knowledge test and the Wordsum. They were then presented with thumbnail images and brief descriptions of the stories they had previously seen, and were told, "Some participants who undertook this survey were shown fake news stories (stories concerning events that did not happen, entirely fabricated by the researchers). If you think you may have been shown any fake stories, please select any story you believe to be fake below". Participants could select as many stories as they wished.

Finally, participants were invited to complete the CRT and then debriefed. Participation took approximately 15 minutes in total.

RESULTS

Descriptive statistics

Leave vs. Remain voters

Of the total valid sample, 311 participants (23.94%) indicated that they had voted Leave in the 2016 referendum, and an additional 48 (3.7%) did not vote but would have voted Leave. A total of 696 participants (53.58%) indicated that they voted Remain, while 167 (12.86%) did not vote but would have voted Remain. A total of 65 participants (5%) reported that they did not vote and were not leaning either way, while 12 participants (0.92%) stated that they could not remember how they voted. In the primary analyses, we collapsed those who voted Leave or Remain with those who said they would have voted the same way. Thus, these analyses include 1,222 participants of whom 29.38% were classed as "Leavers" and 70.62% were classed as "Remainers"; participants who did not express a view on Brexit (those who stated that they didn't vote and weren't leaning either way, or couldn't remember how they voted) were excluded[1]. Leavers and Remainers differed significantly on a number of variables (see Table 1); Remainers were on average younger and better educated, and achieved higher scores on the measures of cognitive ability (Wordsum) and analytical thinking (CRT). Relative to Leavers, Remainers reported themselves to be more liberal, considered the Brexit debate to be more important, and reported higher levels of engagement with the debate (defined as the average of responses to three questions assessing how frequently participants consumed traditional media related to Brexit, engaged with social media content related to Brexit and engaged in discussions with family and friends about Brexit). There was no significant difference in gender distribution or Brexit-related subject knowledge.

[Insert Table 1 here]

[1] The exclusion of these participants slightly reduces the power of the two-way ANOVA described below relative to the target sample size in our preregistration, however the change is negligible. The power of the remaining analyses is not affected by this change.

Response to videos

Participants found the content of the threat video significantly more surprising when it referred to their own group (M = 4.75, SD = 2.22) than when it referred to the other group (M = 3.63, SE = 0.09; t(1222) = 9.51, p < .001, *d* = 0.54). Participants who saw a video referring to their own group also considered the video to be more insulting (M = 6.47, SD = 1.81) than those who saw a video referring to the other side (M = 3.74, SD = 1.78; t(1222 = 26.62, p < .001, *d* = 1.52). Similarly, participants believed the video to be more biased if it described their own group (M = 6.39, SD = 1.67) rather than the other group (M = 4.09, SD = 1.67; t(1222) = 25.60, p < .001, *d* = 1.46). The videos therefore appear to have been successful in their objective of providing insulting, socially threatening content.

True and false memories

In line with our preregistered analysis plan, we considered a story to be remembered if participants responded either "I remember seeing/hearing about this" (hereafter, specific memories) or "I don't remember seeing/hearing about this, but I remember it happening" (hereafter, non-specific memories). On average, participants remembered 27.27% of the false stories (SD = 3.89%), and 49.37% (SD = 22.5%) of the true stories (see Table 2 for statistics for each story). In total, 570 (43.88%) of participants reported a memory for at least one false story, with 138 (10.63%) recalling both false stories. As Table 1 shows, there were no significant differences in the number of true or false stories remembered by Leave and Remain voters. Across both variants of the two fake stories, participants who reported seeing or hearing about the event were more likely to provide a specific source for the memory (e.g. TV, newspaper etc.; 60.57%) than those who reported only remembering the event happening (27.56%), or those who reported merely believing that the event had happened (10.41%). The majority of participants who reported either a specific or non-specific false memory also provided qualitative responses describing their original reaction to the events depicted in the fabricated stories; examination of these responses provides support for the idea that many participants were

remembering specific details about these false events. Some examples are listed below, and the full dataset can be viewed at https://osf.io/rn4ae/.

Election tampering story:

> "Being an ex employee of the council, I thought that the alleged tampering was a bit over the top. Anyone could have put them there" (68-year old male Leave voter, viewed Leave scandals)
>
> "I wasn't surprised–there are always certain individuals who will attempt to tamper with elections" (32-year-old male Leave voter, viewed Remain scandals).
>
> "I felt a bit cheated – what if things had been changed due to this?" (28-year-old female Remain voter, viewed Leave scandals)
>
> "I remember the lack of follow-up coverage offering definite proof of 'tampering'" (53-year-old male Remain voter, viewed Remain scandals)

Wikileaks story:

> "Concerned a little but thought it may be a smear campaign" (45-year-old male Leave voter, viewed Leave scandals)
>
> "I didn't really understand the story anyway at the time. But it was all over the TV and online" (33-year-old female Leave voter, viewed Remain scandals)
>
> "It confirmed in my mind that outside interference in British politics is now deeply entrenched" (54-year-old male Remain voter, viewed Leave scandals)
>
> "Irritated – I believed that this news article was taken out of context and the findings were not accurate" (32-year-old female Remain voter, viewed Remain scandals)

[Insert Table 2 here]

Effects of ideological congruency and identity threat

Our primary preregistered hypotheses in this arm of the study were (1) that participants would form more false memories for Brexit-related fake news stories that are ideologically congruent than for stories that are ideologically incongruent; and (2) that the effect of ideological congruency would be enhanced when a threat was presented to the individuals' sense of group identity as a Leave or Remain supporter, compared with a threat to the opposing group's identity. To evaluate these hypotheses, we collapsed responses across voting groups, and compared those who viewed ideologically congruent stories (i.e. Leavers who viewed scandals about the Remain campaign, and Remainers who viewed scandals about the Leave campaign) with those who viewed ideologically incongruent stories (Leavers or Remainers who viewed scandals about their own side). Similarly, we compared participants who viewed a video that presented a threat to their own group identity (i.e. Leavers who saw the Leave threat video and Remainers who saw the Remain threat video) with those who viewed the video that presented a threat to the other group (i.e. Leavers who saw the Remain threat video and Remainers who saw the Leave threat video)[2]. The proportions of participants who reported a specific memory, non-specific memory, or belief for each story are presented in Figure 2, which collapses the data across the threat conditions. Our preregistered analysis plan defined a false memory as either a specific or non-specific memory of the event, excluding those who merely believed that the event had happened. In the interests of completeness we also describe the pattern of results if false memories are operationalised more narrowly (i.e., if only specific memories are included), or more broadly (e.g., if specific memories, non-specific memories, and beliefs are included).

[Insert Figure 2 here]

[2] Some participants indicated a change in voting preference when asked how they would vote tomorrow compared with how they voted in 2016; 46 participants switched from Leave to Remain, while 36 switched from Remain to Leave. The results reported here did not change significantly if ideological congruency and threat condition were defined based on current (rather than 2016) position on Brexit.

As predicted, a main effect of ideological congruency was observed, such that participants reported more memories for fabricated stories that were congruent with their own ideological position (M = 0.59, SE = 0.03) than for stories that were incongruent (M = 0.50, SE = 0.03; F (1,1218) = 4.67, p = .03, *d* = .11). In other words, participants were more likely to falsely recall scandals that reflected poorly on the opposing side, although the effect was small. An unexpected main effect of threat condition was also observed: participants reported more false memories if they had seen the version of the video that presented a threat to the other group (M = 0.58, SE = 0.03) than if they had seen the version that threatened their own group (M = 0.50, SE = 0.03; F (1,1218) = 4.51, p = .03, *d* = .11); again, this effect was small. Contrary to our hypotheses, there was no significant interaction effect (F (1,1218) = 2.29, p = .13), although a trend was observed in the predicted direction (see Figure 3b). Tests of simple main effects revealed a significant ideology effect among those who viewed the threat to their own group (F(1,585) = 6.70, p = .01, *d* = 0.21), but not among those who viewed the threat to the other group (F(1,607) = 0.03, p = .87, *d* = 0.04).

[Insert Figure 3 here]

Ideological congruency and threat: Exploratory analyses

Additional exploratory analyses revealed the same main effect of ideological congruency regardless of how we defined false memories (specific memory only; specific and non-specific memories; all memories and beliefs). The main effect of threat was not significant in the analysis that included all memories and beliefs, while a borderline effect was observed when analysis was restricted to specific memories only. When non-specific memories and false beliefs were both included in the operationalization of false memories, no significant interaction was found (Figure 3c); however, when we restricted the analysis to specific memories only, this predicted interaction effect was significant (F(1,1218) = 5.25, p = .02; Figure 3a). In this analysis, tests of simple main effects revealed a significant ideology effect among those who viewed the threat to their own group (F(1,585) = 11.52, p < .001, *d*

= 0.28), but not among those who viewed the threat to the other group (F(1,607) = 0.03, p = .87, *d* = 0.01). Full details of these analyses can be found in supplemental materials.

As noted in our preregistration, the hypotheses may only be supported for those who interpret the fabricated stories as intended (i.e., those who consider the fictional scandals to reflect poorly on the campaign in question). Therefore, follow-up logistic regression analyses were conducted for the Wikileaks and election tampering stories, separately for those participants who indicated that the story reflected poorly on the campaign in question (operationalised as a rating greater than 5, where 1 indicated "very well" and 9 indicated "very badly") and for those who did not (a rating less than or equal to 5). In these analyses we defined false memories as either specific or non-specific memories. Further details of these analyses can be found in supplemental materials; briefly, for both the Wikileaks and election tampering stories, the models were significant for those who interpreted the stories as expected, but not for those who interpreted the stories differently. Among those participants who interpreted the stories as expected, effects of ideological congruency were observed for both stories, but the effect of threat condition was only significant for the Wikileaks story. No significant interaction effects were observed in these analyses.

Effects of individual differences in cognitive ability and subject knowledge or engagement

In this arm of the study, our preregistered hypotheses were:

1. That participants with higher scores on measures of cognitive ability and analytical reasoning would form fewer false memories for Brexit-related fake news stories.
2. That participants with higher levels of knowledge, interest and engagement about Brexit would form more false memories for Brexit-related fake news stories.

A multiple regression was performed, in which the outcome variable was the number of false memories reported, and the predictor variables were (1) Brexit subject knowledge test score; (2) subjective importance of referendum outcome; (3) engagement with the Brexit debate; (4) Wordsum score, (5) CRT score, (6) ideological congruency of fabricated stories (where 1 = congruent and 0 = incongruent) and (7) video threat condition (1 = threat to own group, 0 = threat to other group)[3]. In line with our preregistered plan, participants who reported having seen the Wordsum questions or the 7-item CRT before (N = 20) were excluded from analysis. In a deviation from our preregistration, the false memory count variable was modelled using a Poisson distribution rather than a normal distribution as this offered a better fit to the count data, and the assumptions of the analysis were met. Results from this analysis are very similar to the preregistered linear regression, which, in the interests of transparency, is presented in supplemental materials. Several of the continuous predictor variables were moderately correlated with each other, as shown in Table 3, however collinearity values were well within acceptable limits (variance inflation factor < 1.5, tolerance > 0.75 for all predictors). Zero-order correlations for all variables, including the number of true and false memories, can be found in supplemental materials.

[Insert Table 3 here]

The overall model was significant χ^2 (7) = 49.29, p < .001. As Table 4 indicates, the number of false memories reported increased significantly with higher levels of self-reported engagement with the Brexit debate, but there was no significant effect of subject knowledge about Brexit and no effect of

[3] The ideological congruency and threat condition variables were not included in our preregistered plan for this analysis; however in hindsight we believed it important to account for the impact of these variables rather than collapsing across them. The inclusion of these variables does not materially affect the outcome for the other predictor variables, and the new analysis remains adequately powered, with 95% power to detect small effects. The preregistered analysis (containing only the variables labelled 1-5 above) can be found in supplemental materials.

how important participants believed the outcome of the referendum to be. Higher scores on both the Wordsum and CRT were associated with a reduced tendency to report false memories.

[Insert Table 4 here]

Individual differences: Exploratory analyses

A multiple linear regression model was applied to assess the effect of individual differences in subject knowledge, engagement, cognitive ability and reasoning style on the number of true stories for which participants reported a memory. Once more, the overall model was significant ($F(7,1186) = 12.89$, $p < .001$, $R^2 = .07$). Higher scores on measures of subject knowledge and self-reported engagement were associated with more memories for the true stories, but no effect of the subjective importance of the referendum outcome was observed. There was no significant effect of Wordsum score on the number of true memories, but there was a significant effect of CRT score, such that participants who scored higher on analytical reasoning reported fewer true memories (See Table 4). No effects of ideological congruency or threat condition were observed in this analysis.

Comparison of the true and false memory data reveal some interesting patterns. For instance, subject knowledge about Brexit increased memories for true stories but not fabricated stories, while self-reported engagement with the debate increased both true and false memories. Intriguingly, higher analytical reasoning scores were associated with a decrease in reported memories for both true and fabricated stories. One possibility is that increases in reports of true and false memories may be the result of over-claiming on the part of participants who do not wish to admit ignorance – particularly in relation to a topic in which they have expressed considerable interest. Informally interpreting the findings reported thus far could suggest that greater subject knowledge increases sensitivity to the difference between true and false events, whereas engagement and analytical reasoning may simply shift the threshold for reporting a memory. Thus, people who are more engaged with the Brexit

debate may be biased towards 'recognizing' news stories as true, regardless of whether or not they have previously experienced them, while analytical thinkers may apply a stricter criterion when evaluating the stories. Specifically, these participants may have been more alert to the possibility that some of the news stories presented to them were fake, and thus required more evidence before endorsing any story as true.

To test these exploratory hypotheses, we derived d', a measure of signal detection sensitivity, which was computed as the z-score of the hit rate (number of true memories) minus the z-score of the false alarm rate (number of false memories) for each participant. We also derived a measure of response bias ('criterion'), computed as the z-score of the proportion of trials on which the participant reported a memory, for each participant. It is important to note that these measures are derived from just six items – four true stories and two fabricated stories. Evidence suggests that signal detection estimates become more stable as the number of items to be discriminated increases (Pastore & Scheirer, 1974). Thus, the signal detection measures described below should be interpreted cautiously and with reference to the true and false memory count data.

These scores were subjected to the same regression model as above, and the coefficients are reported in Table 4. Both models were significant (d': $F(7,1186) = 8.30$, $p < .001$, $R^2 = .05$; criterion: ($F(7,1186) = 12.93$, $p < .001$, $R^2 = .07$). As expected, greater Brexit-related subject knowledge was significantly associated with sensitivity: participants who were more knowledgeable about Brexit were better able to distinguish between memories for true and false stories, but there was no effect of self-reported engagement on this measure. Both subject knowledge and engagement significantly affected response bias: participants who knew more or were more engaged with the Brexit debate were more likely to report a memory of a story, regardless of whether it was true or false.

CRT scores were negatively associated with response bias, such that participants who were more analytical reported fewer memories; however, analytical reasoning was not significantly associated with sensitivity. Cognitive ability, ideological congruency and identity threat were not significantly associated with response bias or sensitivity.

Identification of fake stories

At the end of the experiment, participants were warned they may have been exposed to fake news and were asked to select any stories they believed were fabricated from a list of all of the stories they had previously seen. Participants identified a mean of 1.70 stories as fake (SD = 1.14, range = 0-6). The Wikileaks story was identified as fake by 479 (36.87%) of participants, and the election tampering story was selected by 696 (53.58%). Selection of the true stories as fake ranged from 4.23% for the story about the falling value of the pound, to 46.57% for the story about Jeremy Corbyn in a Photoshopped chicken suit. Participants were significantly less likely to identify the fabricated stories as fake if they had originally reported remembering the event; for the Wikileaks story, 22.62% of people who originally recalled the event identified the story as fake, compared with 47.44% of those who did not remember it (excluding those who reported a mere belief in the event; χ^2 (1) = 63.02, N = 1090, p < .001, *V* = 0.24). For the election tampering story, 29.33% of those who remembered the event subsequently identified the story as fake, compared with 67.45% of those who did not report remembering the event (χ^2 (1) = 141.71, N = 1152, p < .001, *V* = 0.35).

To determine the factors that affected participants' ability to correctly identify the fake stories, we computed d' as the difference between the standardised hit rate (proportion of fake stories correctly identified as fake) and false alarm rate (proportion of true stories incorrectly identified as fake). These values were entered into a multiple linear regression with the predictors (1) Brexit subject knowledge test score; (2) subjective importance of referendum outcome; (3) engagement with the Brexit debate; (4) Wordsum score, (5) CRT score, (6) ideological congruency of fabricated stories and (7) video threat

condition. As above, participants who reported having seen the CRT and Wordsum questions before were excluded. The model was significant ($F(7, 1186) = 16.50$, $p < .001$, $R^2 = .09$); see Table 5 for regression coefficients. Participants' ability to correctly identify the false stories was significantly impaired by the ideological congruency of the fake stories they had seen, such that participants who viewed stories that reflected poorly on the opposing side were considerably less likely to accurately discriminate between true and fake stories. Identification ability was enhanced among participants with greater Brexit subject knowledge and higher CRT scores, though the effect of CRT score was relatively weak. There was no significant effect of engagement with the Brexit debate, subjective importance of the outcome, Wordsum score or threat condition.

[Insert Table 5 here]

DISCUSSION

This study investigated factors that influence the development of false memories from partisan fake news. Participants each viewed two fabricated stories, one describing election tampering during the 2016 Brexit referendum and another describing a Wikileaks report of illegal campaign donations. Overall rates of false memory were consistent with previous research on rich false memories (Scoboria et al., 2017): each fabricated story was remembered by approximately 22-35% of participants, and 44% of participants reported a false memory for at least one story. Importantly, many participants reported details of their false memories, including the source of the news story and their emotional reaction on first hearing about the fabricated scandals. Approximately 70% of participants in the present sample were classified as "Remainers", a considerably higher rate than the 48% of UK voters who voted to Remain in the 2016 Brexit referendum. This is likely an artefact of the recruitment methods used in the present study; participants on crowdsourcing platforms such as Prolific (used in the present study) or MTurk tend to be younger, better educated and more liberal than the general population (Levay et al., 2016; Peer et al., 2017; Stewart et al., 2017), all factors which differed between Leave and Remain supporters in the present study. This over-representation of Remain supporters is unlikely to have had any impact on our findings however, as our primary analyses collapsed across Leavers and Remainers to evaluate effects of ideological congruency.

As predicted, participants were more likely to falsely recall ideologically congruent fake news. Specifically, Remain voters were more likely to recall fabricated stories implicating the Leave side in wrongdoing, and vice versa; it is however important to note that the magnitude of this effect was smaller than those reported elsewhere (Frenda et al., 2013; Murphy et al., 2019). Previous research has suggested that presenting a threat to social identity increases motivated reasoning and enhances the individual's desire to derogate the opposition (Branscombe et al., 1999). Here we presented participants with a video purporting to be an Australian TV news report, in which either Leave or Remain voters were described in unflattering terms. We expected that viewing the version of the

video that was insulting to the participant's own group would increase false memories about the opposing group. A significant interaction was not however observed in our preregistered analysis, in which false memories were operationalised as either a specific or non-specific memory of the false events. The interaction effect strengthened as false memories were defined more conservatively in exploratory analyses; when analysis was restricted to specific memories only – that is, stories for which participants reported having seen or heard about the event – a modest interaction between ideological congruency and threat condition was observed, such that the effect of ideological congruency was only observed among those who viewed the threat to their own side.

Based on these results from exploratory analysis, we cautiously suggest that when participants viewed a video that insulted their social group, they may have experienced cognitive dissonance as a result of the conflict between their positive view of their in-group and the negative assessment presented in the video. Cognitive dissonance is often described a state of psychological discomfort caused by holding conflicting cognitions, but it is also understood to be a motivational state that prompts the individual to alleviate the discomfort by removing or changing one of the conflicting cognitions, or by adding a new cognition to resolve the conflict (Elliot & Devine, 1994; Festinger, 1957). When faced with information that is threatening to one's sense of self, this dissonance may be resolved by creating new cognitions that justify the behaviour of the in-group or derogate members of the out-group (Gawronski, 2012). Resolving the dissonance can sometimes result in memory distortion, such as when participants who are forced to espouse attitudes they disagree with later misremember their initial attitudes as being closer to the position they have just articulated (Rodriguez & Strange, 2015). Thus, we propose that participants in the present experiment were driven to resolve the cognitive dissonance created by the threatening video by preferentially recalling fake news stories that derogated the opposing campaign; the same cognitive process may have suppressed the construction of memories for events that reflected poorly on the participant's own side. We speculate that similar

effects might also be observed for true memories, with preferential recall of ideologically congruent events, though the present study was not designed to test this hypothesis.

The interaction between ideological congruency and video threat condition was small, and only emerged as statistically significant when we applied the most conservative definition of a false memory. When the definition of memory included non-specific memories (as planned in our preregistration) or false beliefs, no significant interactions were observed, although a clear effect of ideological congruency remained. While false memories arise through a relatively effortful constructive process, often as a result of source confusion (Hyman & Kleinknecht, 1999; Mazzoni et al., 2001; Strange et al., 2005), false beliefs may come about through a simple evaluation of the plausibility of the story (Mazzoni & Kirsch, 2002). Indeed, Scoboria et al. (2004) suggest that, under most circumstances, plausibility is a prerequisite for belief in the occurrence of an event, which is itself a prerequisite for memory. We therefore speculate that forming a false belief does not require the same cognitive effort as forming a false memory; thus, the effects of ideological congruency were observed even in the absence of the additional cognitive dissonance provided by the threatening video. By the same token, the extra cognitive effort required to form a specific, detailed memory may explain why the effects were stronger when analysis was restricted to specific memories only. This explanation is purely speculative, and further confirmatory research will be required to support it, perhaps by explicitly manipulating the level of specificity with which participants are asked to recall fabricated events. Future research might also examine whether more consistent effects are observed following stronger manipulations of identity threat. Our manipulation checks demonstrated that participants found the video about their own group to be surprising, insulting and biased, but it is possible that even subtle alterations – for example, including subtitles to reinforce the threatening message – would induce a more polarising effect.

Our second research question concerned the role of individual differences in cognitive ability and analytical reasoning, as well as participants' knowledge, interest and engagement in the Brexit debate. We discuss the cognitive variables later in this section; considering the knowledge and engagement variables, our initial preregistered analysis indicated that the number of false memories was positively associated with self-reported engagement in the Brexit debate, defined as engagement with traditional media, social media and conversations with friends and family on the topic. Interestingly, objectively assessed knowledge about Brexit did not affect false memory formation, in contrast with previous reports relating expertise to increased false memories (Baird, 2003; Castel et al., 2007; Mehta et al., 2011); however, the data supported prior evidence that self-reported interest in a topic increases the rate of false memories (O'Connell & Greene, 2017). A source monitoring interpretation of these data (Johnson et al., 1993; Mitchell & Johnson, 2000) might suggest that a more developed schema, arising from increased knowledge or frequent engagement with a topic, creates multiple overlapping memory traces which may be activated by a novel (but related) story, triggering a sense of familiarity, and thence a constructed memory of the event. This account might predict that the false memory rate would be positively correlated with both objectively assessed knowledge and self-reported engagement, in contrast with the dissociation observed here. In interpreting the effects of engagement with the topic, it may be better to consider the possibility that participants are simply overclaiming, or imposing a very low threshold for reporting a memory, in line with previous evidence that fake news acceptance is associated with a general tendency to overclaim one's knowledge (Pennycook & Rand, 2018).

To further investigate this question, we conducted exploratory signal detection analysis. Brexit subject knowledge was significantly associated with response bias, such that more knowledgeable participants were more likely to respond that they remembered any given story. Subject knowledge also increased sensitivity, suggesting that knowing more about the topic led participants to report more memories for the true stories while being somewhat less susceptible to false memories for the

fabricated stories. Interestingly, self-reported engagement with the Brexit debate was not associated with sensitivity, but was associated with response bias. Thus, participants who were more interested and engaged in the topic were more likely to report a memory for any given story, but were not better able to discriminate between memories for true and false stories. These data lend support to the overclaiming hypothesis described above, suggesting that the likelihood of forming a false memory in response to a fake news story is influenced by the individual's desire to seem well-informed, rather than by their actual level of knowledge (cf. Mehta et al., 2011). In this study, subject knowledge was assessed via an objective test, while engagement was self-reported; given the highly topical nature of the Brexit debate, and the overclaiming issues described above, it is possible that participants who wished to present themselves as knowledgeable and engaged may have over-reported their actual engagement with media and discussions on the subject. In further investigating the contribution of subject knowledge, interest and engagement, future researchers may therefore wish to obtain an objective measure of engagement to tease apart effects of the self-report nature of the measure from the underlying construct being assessed.

Higher scores on measures of cognitive ability and analytical reasoning were associated with a reduced tendency to report false memories. This finding appears to support the suggestion that failures of analytical reasoning underlie the perceived truthfulness of fake news (Pennycook & Rand, 2019), and that cognitive ability is an important factor in the generation of false memories (Murphy et al., 2019). These results are also in line with research identifying cognitive ability as a critical factor in eyewitness susceptibility to misinformation (Greene et al., 2020; Zhu et al., 2010). Cognitive ability was assessed using the Wordsum, which is highly correlated with full-scale IQ, but it is important to note that there are some limitations to employing a vocabulary-based measure of cognitive ability. First, the Wordsum assesses crystallised rather than fluid intelligence (Malhotra et al., 2007) and is therefore partially confounded with education and socioeconomic status. Second, there is the possibility of underestimating the cognitive ability of individuals whose English-language skills do not match their

general level of intelligence. This may be the case for people who speak English as a second language, (although approximately 97% of our participants reported English as their first language), or for people with a specific language impairment or learning disability (e.g. dyslexia).

Once more, we can look to the signal detection measures for a deeper understanding of the processes underlying effects of cognitive ability and analytical reasoning. General cognitive ability did not have a significant effect on either signal detection measure. In contrast, analytical reasoning (measured using the CRT) was associated with response bias, such that more analytical respondents were less likely to report a memory for any given story, but it was not significantly associated with the differential reporting of true vs. false memories (i.e. sensitivity). Thus, it appears that analytical reasoning increased participants' overall suspicion towards the stories, and therefore decreased false memories, but did not make participants better at discriminating between memories for true and false stories.

In applying a signal detection approach to this study of false memories, we recognise that the present experiment differs from a traditional recognition memory paradigm in which the experimenters have control over participants' previous exposure to stimuli (e.g. Lockhart & Murdock, 1970; Morrell et al., 2002; Wixted, 2007). Here, participants may or may not have encountered the true stories in their day-to-day lives, adding uncertainty to the calculation of hits (since participants may correctly report that they do not remember a true story if they have simply never encountered it). Moreover, as noted in the Results section, the signal detection analyses rely on responses to a small number of items, and may thus be relatively unstable (Pastore & Scheirer, 1974). These data should therefore be interpreted with caution. Despite these caveats, signal detection measures provide valuable information about participants' interaction with the stimuli used in this experiment, allowing us to separate an individual's threshold for 'remembering' a story from their ability to discriminate true from false. This is particularly important as raw estimates of misinformation acceptance (whether measured as false

memories or false beliefs) may be biased by a respondent's overall degree of scepticism or credulity, a trait sometimes referred to as 'reflexive open-mindedness' (Pennycook et al., 2015; Pennycook & Rand, 2018).

The false memories reported in this study were remarkably resistant to post-exposure warnings about fake news; participants who initially reported a memory of a fabricated story were less than half as likely to identify that story as fake when asked to select the fabricated stories from a list. The effects of ideological congruency were similarly persistent; participants were considerably worse at correctly identifying fake stories during debriefing if the fabricated stories they had seen reflected poorly on their out-group. This replicates our previous findings in a similar study (Murphy et al., 2019), and underscores concerns about the effectiveness of post-exposure fact-checking (Young et al., 2018). As with initial reporting of false memories, better Brexit subject knowledge and higher levels of analytical reasoning increased identification accuracy, suggesting that these factors are beneficial regardless of whether participants have been alerted to the presence of fake news.

The data presented in this paper demonstrate the ease with which false memories can be formed from fabricated news stories, and emphasise the importance of both social and personal factors in determining how and when an individual will fall for fake news. As noted above, UK citizens are regularly subjected to social identity threats in the form of media coverage that derogates either Leavers or Remainers. The same is true in other polarised political contexts, where people frequently encounter insulting content about groups with which they identify. Our study provides suggestive evidence that exposure to a social identity threat may enhance motivated reasoning and increase the tendency to form false memories for events that reflect poorly on the out-group while protecting the in-group. Because false memories can have significant consequences for real-world behaviour (Laney & Loftus, 2017), this has implications both for the historical recollection of political events and for the integrity of political contests and elections. To provide a recent example, the latter half of 2020 saw

protests across the United States and around the world in response to police brutality against Black people and other people of colour. Very rapidly, existing social and political identities coalesced around distinct media narratives, supporting either the protestors or the police, and misinformation spread through social media. The data presented in this paper suggest that observers' memories of these kinds of political events are likely to be shaped along ideological lines, leading to conflicting sets of recollections with implications for downstream behaviour. Perhaps the most important message to take from these data is that people are more likely to form false memories in response to fake news that reflects poorly on an opposing group, and by extension, reflect well on their own group. This might be especially true when people's group identity feels threatened. Importantly, our data also show that individual differences in a range of factors predict susceptibility to misinformation, demonstrating that group membership is not the sole determinant of who will fall for partisan fake news.

FUNDING DETAILS

This work was not supported by any research grants.

DISCLOSURE STATEMENT

The authors have no potential conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

The data and materials associated with this paper may be found at https://osf.io/rn4ae/.

REFERENCES

Axe-Browne, A., & Hansen, M. E. (2020). Still dividing the electorate? Brexit and voter evaluation of candidates. *Journal of Elections, Public Opinion and Parties*, 1-23.

Baird, R. R. (2003, //). Experts sometimes show more false recall than novices: a cost of knowing too much. *Learning and Individual Differences, 13*(4), 349-355. https://doi.org/http://dx.doi.org/10.1016/S1041-6080(03)00018-9

BBC News. (2016). *EU Referendum Results*. Retrieved 25th May from https://www.bbc.com/news/politics/eu_referendum/results

Bialek, M., & Pennycook, G. (2018). The cognitive reflection test is robust to multiple exposures. *Behavior Research Methods, 50*(5), 1953-1959.

Bolsen, T., Druckman, J. N., & Cook, F. L. (2014). The influence of partisan motivated reasoning on public opinion. *Political Behavior, 36*(2), 235-262.

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. *Social identity: Context, commitment, content*, 35-58.

Branscombe, N. R., & Wann, D. L. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European Journal of Social Psychology, 24*(6), 641-657.

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition, 8*(1), 108-117.

Castel, A. D., McCabe, D. P., Roediger, H. L., & Heitman, J. L. (2007, January 1, 2007). The Dark Side of Expertise: Domain-Specific Memory Errors. *Psychological science, 18*(1), 3-5. https://doi.org/10.1111/j.1467-9280.2007.01838.x

Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of personality and social psychology, 85*(5), 808.

Devine, C. J. (2015). Ideological social identity: Psychological attachment to ideological in-groups as a political phenomenon and a behavioral influence. *Political Behavior, 37*(3), 509-535.

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change, 9*(2), 111-119.

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review, 107*(1), 57-79.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of personality and social psychology, 67*(3), 382.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford university press.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*(4), 25-42.

Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology, 49*(2), 280-286.

Gawronski, B. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social cognition, 30*(6), 652-668.

Grant, P. R., & Brown, R. (1995). From ethnocentrism to collective protest: Responses to relative deprivation and threats to social identity. *Social Psychology Quarterly*, 195-212.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1).

Greene, C. M., Maloney-Derham, R., & Mulligan, K. (2020). Effects of perceptual load on eyewitness memory are moderated by individual differences in cognitive ability. *Memory, 28*(4), 450-460. https://doi.org/https://doi.org/10.1080/09658211.2020.1729811

Greene, C. M., & Murphy, G. (2020). Can fake news really change behaviour? Evidence from a study of COVID-19 misinformation. *Under review*. https://doi.org/10.31234/osf.io/qfnm3

Hart, P. S., & Nisbet, E. C. (2012). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication research, 39*(6), 701-723.

Hinde, S. (2017). Brexit and the media. *Hermès, La Revue*(1), 80-86.

Hobolt, S., Leeper, T. J., & Tilley, J. (2020). Divided by the vote: affective polarization in the wake of the Brexit referendum. *British Journal of Political Science*.

Hoppe, E. I., & Kusterer, D. J. (2011). Behavioral biases and cognitive reflection. *Economics Letters, 110*(2), 97-100.

Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and personality psychology compass, 2*(1), 204-222.

Huang, M.-H., & Hauser, R. M. (1996). *Trends in black-white test score differentials: II. The WORDSUM vocabulary test*. Center for Demography and Ecology, University of Wisconsin--Madison.

Hyman, I. E., & Kleinknecht, E. E. (1999). False childhood memories: Research, theory, and applications. In L. M. Williams & V. L. Banyard (Eds.), *Trauma and memory* (pp. 175–188). Sage, Thousand Oaks.

Jaschinski, U., & Wentura, D. (2002). Misleading postevent information and working memory capacity: An individual differences approach to eyewitness memory. *Applied Cognitive Psychology, 16*(2), 223-231.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993, Jul). Source monitoring. *Psychological Bulletin, 114*(1), 3-28.

Kahan, D. M. (2015). The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*, 1-16.

Kahan, D. M. (2017). Misconceptions, misinformation, and the logic of identity-protective cognition.

Laney, C., & Loftus, E. F. (2017). False memories matter: the repercussions that follow the development of false memory. In R. A. Nash & J. Ost (Eds.), *False and distorted memories* (pp. 153-165). Psychology Press.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., & Rothschild, D. (2018). The science of fake news. *Science, 359*(6380), 1094-1096.

Lee, K. (2004). Age, neuropsychological, and social cognitive measures as predictors of individual differences in susceptibility to the misinformation effect. *Applied Cognitive Psychology, 18*(8), 997-1019.

Levay, K. E., Freese, J., & Druckman, J. N. (2016, 2016/01/01). The Demographic and Political Composition of Mechanical Turk Samples. *SAGE Open, 6*(1), 2158244016636433. https://doi.org/10.1177/2158244016636433

Lindsay, D. S. (1994). Memory source monitoring and eyewitness testimony. *Adult eyewitness testimony: Current trends and developments*, 27-55.

Lockhart, R. S., & Murdock, B. B. (1970). Memory and the theory of signal detection. *Psychological Bulletin, 74*(2), 100.

Malhotra, N., Krosnick, J. A., & Haertel, E. (2007). The psychometric properties of the GSS Wordsum vocabulary test. *GSS Methodological Report, 11*.

Mazzoni, G., & Kirsch, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model.

Mazzoni, G. A., Loftus, E. F., & Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: A little plausibility goes a long way. *Journal of Experimental Psychology: Applied, 7*(1), 51.

Mehta, R., Hoegg, J., & Chakravarti, A. (2011). Knowing too much: Expertise-induced false recall effects in product comparison. *Journal of Consumer Research, 38*(3), 535-554.

Meisenberg, G. (2015). Verbal ability as a predictor of political preferences in the United States, 1974–2012. *Intelligence, 50*, 135-143.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. *The Oxford handbook of memory*, 179-195.

Morrell, H. E., Gaitan, S., & Wixted, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of experimental psychology: Learning, Memory, and Cognition, 28*(6), 1095.

Murphy, G., Loftus, E. F., Grady, R. H., Levine, L. J., & Greene, C. M. (2019). False Memories for Fake News During Ireland's Abortion Referendum. *Psychological science, 30*(10), 1449-1459.

Nash, R. A. (2018). Changing beliefs about past public events with believable and unbelievable doctored photographs. *Memory, 26*(4), 439-450.

O'Connell, A., & Greene, C. M. (2017). Not strange but not true: self-reported interest in a topic increases false memory. *Memory, 25*(8), 969-977.

Pastore, R., & Scheirer, C. (1974). Signal detection theory: Considerations for general application. *Psychological Bulletin, 81*(12), 945.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017, 2017/05/01/). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153-163. https://doi.org/https://doi.org/10.1016/j.jesp.2017.01.006

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision making, 10*(6), 549-563.

Pennycook, G., & Rand, D. G. (2018). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39-50.

Rodriguez, D. N., & Strange, D. (2015). False memories for dissonance inducing events. *Memory, 23*(2), 203-212.

Sacchi, D. L. M., Agnoli, F., & Loftus, E. F. (2007). Changing history: Doctored photographs affect memory for past public events [Article]. *Applied Cognitive Psychology, 21*(8), 1005-1022. https://doi.org/10.1002/acp.1394

Schaller, M. (1992). In-group favoritism and statistical reasoning in social inference: Implications for formation and maintenance of group stereotypes. *Journal of personality and social psychology, 63*(1), 61.

Scoboria, A., Lynn, S. J., Hessen, J., & Fisico, S. (2007). So that's why I don't remember: Normalising forgetting of childhood events influences false autobiographical beliefs but not memories. *Memory, 15*(8), 801-813.

Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology, 18*(7), 791-807. https://doi.org/https://doi.org/10.1002/acp.1062

Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., & Hyman, I. E. (2017). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory, 25*(2), 146-163.

Stewart, N., Chandler, J., & Paolacci, G. (2017, 2017/10/01/). Crowdsourcing Samples in Cognitive Science. *Trends in Cognitive Sciences, 21*(10), 736-748. https://doi.org/https://doi.org/10.1016/j.tics.2017.06.007

Strange, D., Garry, M., Bernstein, D. M., & Lindsay, D. S. (2011). Photographs cause false memories for the news. *Acta Psychologica, 136*(1), 90-94.

Strange, D., Gerrie, M. P., & Garry, M. (2005). A few seemingly harmless routes to a false memory. *Cognitive Processing, 6*(4), 237-242.

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Brooks/Cole.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision making, 11*(1), 99.

Thorndike, R. L., & Gallup, G. H. (1944). Verbal intelligence of the American adult. *The Journal of General Psychology, 30*(1), 75-85.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*(2), 147-168.

Van Bavel, J. J., & Pereira, A. (2018, 2018/03/01/). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences, 22*(3), 213-224. https://doi.org/https://doi.org/10.1016/j.tics.2018.01.004

Voci, A. (2006). The link between identification and in-group favouritism: Effects of threat to social identity and trust-related emotions. *British Journal of Social Psychology, 45*(2), 265-284.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological review, 114*(1), 152.

Young, D. G., Jamieson, K. H., Poulsen, S., & Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck. org and FlackCheck. org. *Journalism & Mass Communication Quarterly, 95*(1), 49-75.

Zhu, B., Chen, C., Loftus, E. F., Lin, C., He, Q., Chen, C., Li, H., Xue, G., Lu, Z., & Dong, Q. (2010). Individual differences in false memory from misinformation: Cognitive factors. *Memory, 18*(5), 543-555.

TABLES

Table 1. Comparison of demographic and cognitive variables between Leavers and Remainers

	Leavers	**Remainers**	**Comparison**
	M (SD)	**M (SD)**	
Age	39.98 (13.20)	34.64 (11.89)	$t(1220) = 6.91$, $p < .001$, $d = 0.43$
Gender			χ^2 (2) = 0.84, $N = 1222$, $p = 0.66$, $V = 0.03$
Male	N = 130 (36.21%)	N = 296 (34.30%)	
Female	N = 228 (63.51%)	N = 562 (65.12%)	
Other	N = 1 (0.28%)	N = 5 (0.58%)	
Education			$U(N_{Leave} = 359, N_{Remain} = 863) = 107745.5$, $p <$
No formal education	N = 2 (0.56%)	N = 1 (0.12%)	.001, $r_{biserial} = -0.30$
Primary school	N = 1 (0.28%)	N = 1 (0.12%)	
GCSEs/O-Levels	N = 78 (21.73%)	N = 64 (7.42%)	
A-levels	N = 62 (17.27%)	N = 114 (13.21%)	
Some college/university	N = 86 (23.96%)	N = 160 (18.54%)	
Undergraduate degree	N = 102 (28.41%)	N = 363 (42.06%)	
Postgraduate degree	N = 25 (6.96%)	N = 142 (16.45%)	
Doctoral degree	N = 3 (0.84%)	N = 18 (2.09%)	
Liberal-Conservative	5.46 (1.58)	3.50 (1.54)	$t(1220) = 20.06$, $p < .001$, $d = 1.26$
Importance of Brexit outcome	6.60 (1.80)	7.13 (1.63)	$t(1220) = -4.95$, $p < .001$, $d = -0.31$

Engagement (media/social media/family & friends)	3.20 (0.88)	3.48 (0.82)	$t(1220) = -5.32$, $p < .001$, $d = -0.33$
Brexit subject knowledge	6.86 (2.12)	7.02 (2.08)	$t(1219) = -1.22$, $p = .22$, $d = -0.08$
Wordsum score	6.85 (1.78)	7.27 (1.77)	$t(1218 = -3.76$, $p < .001$, $d = -0.24$
CRT Score	2.94 (1.92)	3.43 (1.94)	$t(1215) = -4.02$, $p < .001$, $d = -0.25$
True stories remembered	1.92 (1.12)	2.03 (1.02)	$t(1220) = -1.64$, $p = .10$, $d = -0.10$
False stories remembered	0.55 (0.69)	0.54 (0.67)	$t(1220) = 0.06$, $p = .95$, $d = 0.004$

Table 2. Responses to fabricated and true stories: Number and percentage of participants who endorsed each response option

	Fabricated stories			
Response	**Wikileaks: Leave**	**Wikileaks: Remain**	**Election tampering: Leave**	**Election tampering: Remain**
I remember seeing/hearing about this	147 (22.83%)	100 (15.27%)	139 (21.58%)	115 (17.56%)
I don't remember seeing/hearing about this, but I remember it happening	64 (9.94%)	56 (8.55%)	35 (5.43%)	52 (7.94%)
I don't remember this, but I believe it happened	109 (16.93%)	100 (15.27%)	72 (11.18%)	75 (11.45%)
I remember this differently	22 (3.42%)	29 (4.43%)	9 (1.4%)	10 (1.53%)
I don't remember this	302 (46.89%)	370 (56.49%)	389 (60.4%)	403 (61.53%)
Total N	*644*	*655*	*644*	*655*

	True stories			
	Jeremy Corbyn	**Value of pound**	**Mike Pence**	**Ford factory**
I remember seeing/hearing about this	428 (32.95%)	913 (70.28%)	208 (16.01%)	615 (47.34%)
I don't remember seeing/hearing about this, but I remember it happening	87 (6.7%)	91 (7.01%)	114 (8.78%)	109 (8.39%)
I don't remember this, but I believe it happened	135 (10.39%)	133 (10.24%)	225 (17.32%)	218 (16.78%)
I remember this differently	54 (4.16%)	32 (2.46%)	17 (1.31%)	47 (3.62%)
I don't remember this	595 (45.8%)	130 (10.01%)	735 (56.58%)	310 (23.86%)
Total N	*1299*	*1299*	*1299*	*1299*

Table 3. Correlation coefficients for individual differences measures

	Brexit subject knowledge	Engagement	Importance of outcome	Wordsum score	CRT score
Brexit subject knowledge	—				
Engagement	0.22**	—			
Importance of outcome	0.25	0.4**	—		
Wordsum score	0.42**	0.18**	0.19**	—	
CRT score	0.35**	0.05	0.07*	0.36**	—

*p < .05

**p < .001

Table 4. Regression coefficients for effects of individual differences measures on false memory count, true memory count, sensitivity (d') and response bias (criterion), controlling for ideological congruency and threat condition.

						95% CI (B)	
Predictor	**B**	**SE (B)**	**$\beta^{\dagger}$**	**Wald χ^2**	**p**	**Lower**	**Upper**
False memory count[a]							
(Intercept)	-0.66	0.25	0.52	7.08	.008	-1.14	-0.17
Brexit subject knowledge	-0.01	0.02	0.99	0.19	.66	-0.05	0.03
Engagement*	0.21	0.05	1.23	16.58	< .001	0.11	0.31
Importance of outcome	-0.003	0.031	0.99	0.02	.90	-0.05	0.05
Wordsum score*	-0.05	0.02	0.95	4.87	.03	-0.10	-0.006
CRT score*	-0.07	0.02	0.93	9.57	.002	-0.11	-0.025
Ideological congruency	-0.12	0.08	0.88	2.46	.12	-0.28	0.03
Threat condition	0.16	0.08	1.17	3.81	.051	-0.001	0.31
True memory count[b]							
(Intercept)	0.82	0.18		4.52	< .001	0.46	1.17
Brexit subject knowledge*	0.08	0.02	0.17	5.23	< .001	0.05	0.12
Engagement*	0.21	0.04	0.17	5.54	< .001	0.14	0.29
Importance of outcome	0.02	0.02	0.04	1.28	0.20	-0.01	0.06
Wordsum score	-0.02	0.02	-0.03	-0.81	0.42	-0.05	0.02
CRT score*	-0.05	0.02	-0.09	-2.76	0.01	-0.08	-0.01
Ideological congruency	-0.03	0.06	-0.02	-0.55	0.58	-0.15	0.08
Threat condition	.005	0.06	0.003	0.09	0.93	-0.11	0.12
Sensitivity (d')[b]							
(Intercept)	-1.12	0.21		-5.32	< .001	-1.54	-0.71
Brexit subject knowledge*	0.09	0.02	0.15	4.69	< .001	0.05	0.13
Engagement	0.03	0.04	0.02	0.72	0.47	-0.06	0.12
Importance of outcome	0.03	0.02	0.04	1.13	0.26	-0.02	0.07
Wordsum score	0.03	0.02	0.04	1.35	0.18	-0.01	0.07
CRT score	0.01	0.02	0.02	0.60	0.55	-0.03	0.05
Ideological congruency	-0.13	0.07	-0.05	-1.89	0.06	-0.27	-0.007

Threat condition	0.13	0.07	0.05	1.86	0.06	-0.007	0.005
Response bias (criterion)[b]							
(Intercept)	-0.82	0.17		-4.79	< .001	-1.15	-0.48
Brexit subject knowledge*	0.06	0.02	0.12	3.68	< .001	0.03	0.09
Engagement*	0.23	0.04	0.20	6.46	< .001	0.16	0.30
Importance of outcome	0.02	0.02	0.03	0.91	0.36	-0.02	0.05
Wordsum score	-0.03	0.02	-0.06	-1.82	0.07	-0.07	0.002
CRT score*	-0.06	0.02	-0.12	-3.80	< .001	-0.09	-0.03
Ideological congruency	0.03	0.06	0.01	0.45	0.65	-0.08	0.13
Threat condition	-0.06	0.06	-0.03	-1.00	0.32	-0.17	0.05

* Significant predictor ($p < .05$)

[a] Poisson regression

[b] Linear regression

[†]For Poisson regressions (i.e. false memory count), β (Exp (B)) is given as 1 for no effect, with values > 1 for positive effects and < 1 for negative effects. For linear regressions (i.e. true memory count), β is given as 0 for no effect with values < 0 for negative effects and > 0 for positive effects.

Table 5. Regression coefficients for discrimination between true and fake stories (d' values) when asked to identify the fake stories from a list

						95% CI	
	B	**S.E. (B)**	**β**	**t**	**p**	**Lower**	**Upper**
(Intercept)	-1.26	0.23		-5.43	< .001	-1.72	-0.81
Brexit subject knowledge	0.12	0.02	0.19	5.85	< .001	0.08	0.16
Engagement	0.02	0.05	0.01	0.41	0.68	-0.08	0.12
Importance of outcome	0.01	0.02	0.02	0.61	0.54	-0.03	0.06
Wordsum score	0.04	0.02	0.05	1.66	0.10	-0.007	0.09
CRT score	0.04	0.02	0.06	2.00	0.05	0.001	0.08
Ideological congruency	-0.43	0.08	-0.16	-5.60	< .001	-0.57	-0.28
Threat condition	0.07	0.08	0.02	0.87	0.38	-0.08	0.21

FIGURES

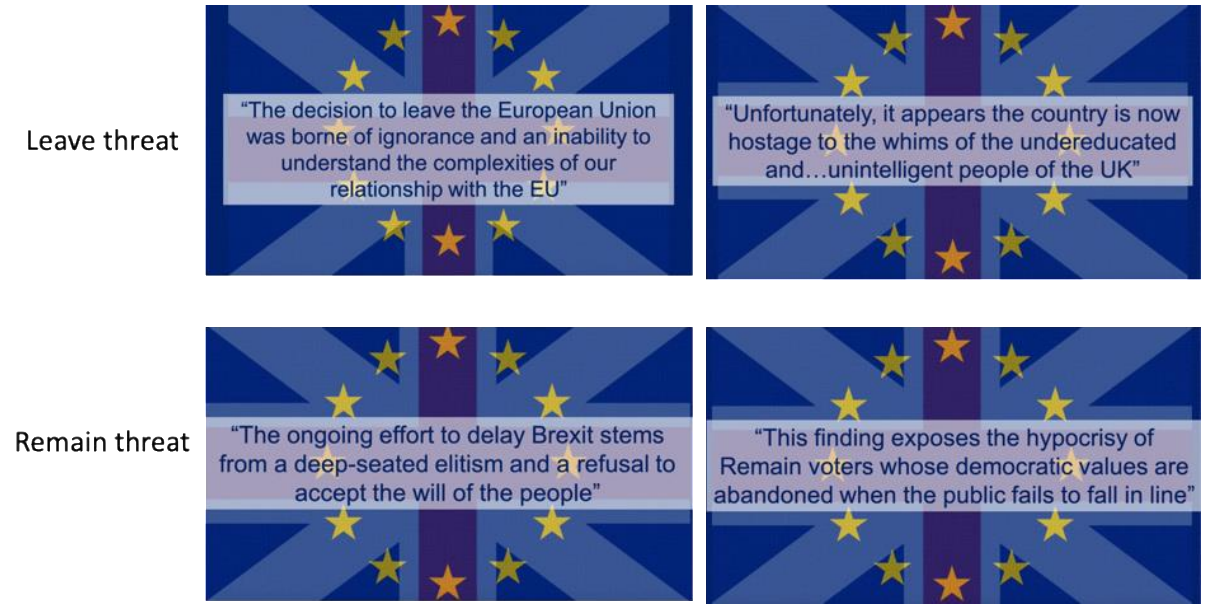

Figure 1. Inflammatory quotes from videos designed to present an identity threat to Leave voters (top row) and Remain voters (bottom row).

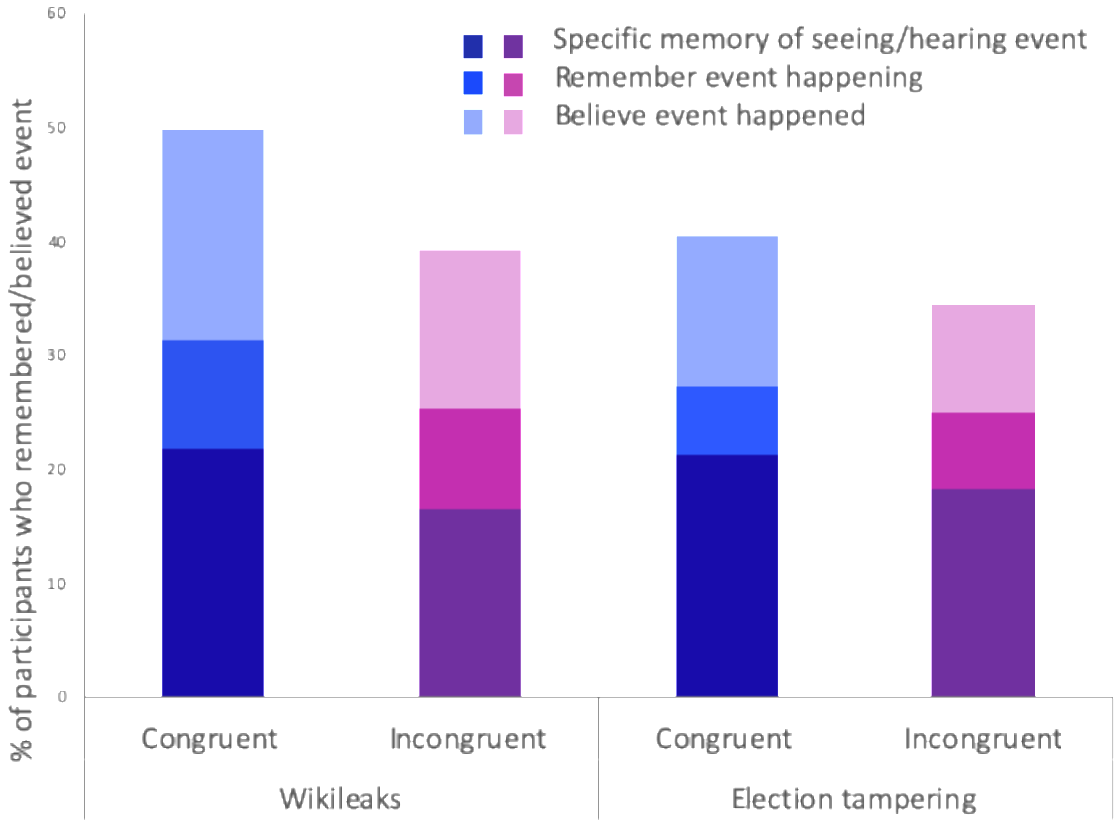

Figure 2. Proportion of participants who reported a specific memory, non-specific memory, or belief for ideologically congruent and incongruent stories

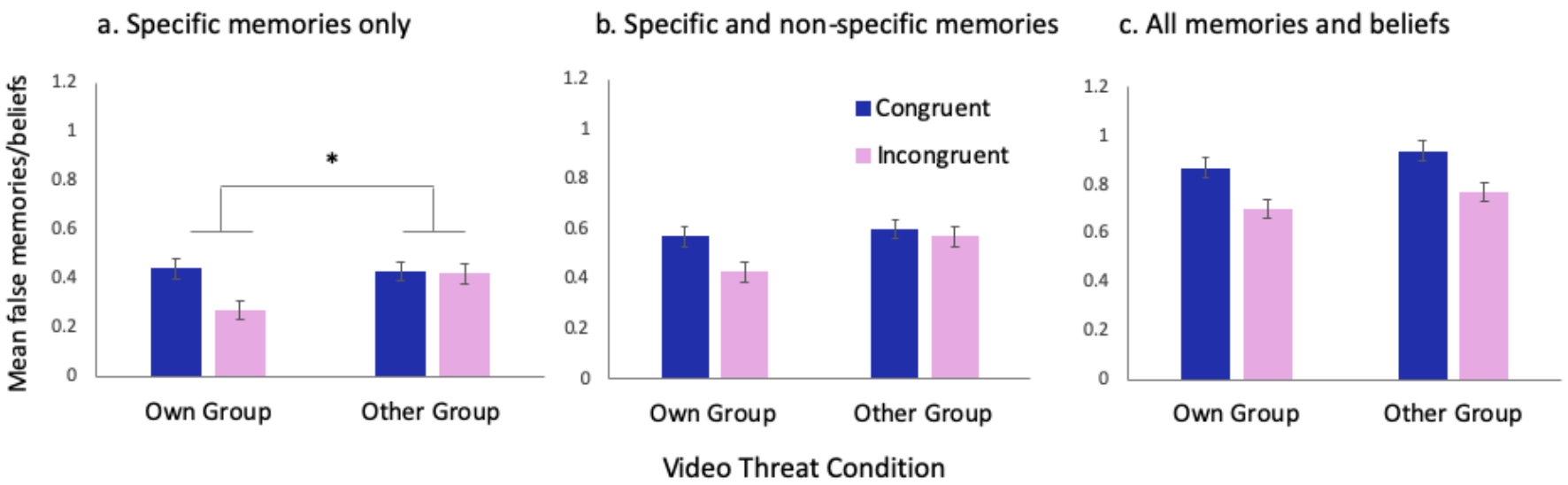

Figure 3. Number of false memories or beliefs reported for fabricated news stories as a function of ideological congruency and threat condition. (a) Specific memories only; (b) specific and non-specific memories, per our preregistered plan; (c) all memories and beliefs. Error bars represent standard error of the mean.

*$p < .05$

FIGURE LEGENDS

Figure 1. Inflammatory quotes from videos designed to present an identity threat to Leave voters (top row) and Remain voters (bottom row).

Figure 2. Proportion of participants who reported a specific memory, non-specific memory, or belief for ideologically congruent and incongruent stories

Figure 3. Number of false memories or beliefs reported for fabricated news stories as a function of ideological congruency and threat condition. (a) Specific memories only; (b) specific and non-specific memories, per our preregistered plan; (c) all memories and beliefs. Error bars represent standard error of the mean. *$p < .05$