

A post-truth public? How the impact of fact-checks is undermined by post-truth rhetoric

Christine Stedtnitz, ^{*} John Bartle, [†] Robert Johns, [‡]

Abstract

While fact-checks go some way to increasing the accuracy of public beliefs, they meet with considerable resistance. The motivations behind that resistance are much clearer than the mechanisms through which ostensibly neutral and factual information is rejected. In a survey experiment we demonstrate how various strands of what we term ‘post-truth’ rhetoric can immediately undermine the effects of fact-checks. We first assessed perceptions about immigration and exposed those with misperceptions to a fact-check. Then we exposed treatment groups to one of three variants of post-truth rhetoric, not directly refuting the accuracy of the fact-check but suggesting that people should take the evidence ‘with a big pinch of salt’. This rhetoric undid almost one third of the positive effect of the fact check on the accuracy of respondents’ beliefs. The effect was significant and similar in size regardless of the variant of rhetoric used or the credentials of the commentator.

Keywords: fact-checking, misperceptions, resistance to correction, post-truth

^{*}London School of Economics and Political Science, c.stedtnitz@lse.ac.uk

[†]University of Essex, jbartl@essex.ac.uk

[‡]University of Essex, rajohn@essex.ac.uk

Introduction

An expert or academic has more chance of leading readers astray with their biased and cherry picked information which is the same as lying.

Madison [@MadisonTX766], April 22nd, 2023, Twitter.

Trust your experience and personal research over anything else. Especially over hacks like politifacts. #walkaway.

Brian Sanders [@Brian_Sanders3], August 15th, 2019, Twitter.

Feelings aren't facts but trust your intuition regardless. If it feels off, troubling or weird then there's a high chance it's not right for you.

Master Apprentice [@truekingmaker], January 12th, 2023, Twitter.

Public debates are complex and characterised by a bewildering series of claims and counterclaims. This complexity has increased in recent years. Before the dawn of the internet, parties, experts, and the media used to structure the national conversation. In the last twenty years or so, however, social media have provided almost everyone with access to information and platforms on which to express their views. Experts are present on these platforms but they are outnumbered and often outgunned in a noisy and confusing public debate, and their assertions of the facts are routinely greeted by the kinds of replies that opened this article. Such responses and their popularity – the three tweets between them have over 250 likes – well illustrate Francis Fukuyama's (2017) definition of a 'post-fact' world as one in which 'virtually all authoritative information sources' are 'called into question and challenged by contrary facts of dubious quality and provenance'. However meagre their own credentials, those challenging expert information in this way are potentially persuasive either because they are recognised as political allies or because they are personally trusted and it is easier for most of us to abandon facts than friendships. The key question for this article is: how persuasive are such 'post-truth' rhetorical challenges to authoritative expert evidence?

Journalists, public bodies, charities, and academics have responded to the concerns about a

‘post-fact’ or ‘post-truth’ world. In particular, there are numerous fact-checking organisations seeking to correct misinformation. In the US, prominent such services include FactCheck.org, PolitiFact and Snopes. In the UK, the empirical focus for this study, the leading public service broadcasters established fact-checks in the mid-2000s (Graves and Cherubini 2016), and Full Fact, a registered charity founded in 2009 to counter misinformation, has partnered with Facebook to provide fact-checking services on that platform since 2019. Many other countries have their own institutions, and the International Fact-Checking Network was established in 2015 to strengthen the sector.

A host of studies have shown that such attempts to correct misinformation enjoy some but only limited success. Three meta-analyses show that people’s perceptions usually shift in the direction of the fact-check but often not by much (Chan et al. 2017; Walter and Murphy 2018; Walter et al. 2020). The reason why this success is often limited is also well established: people are motivated reasoners, susceptible to confirmation bias (Kunda 1990). We all have cognitive, ideological and group-based motivations to be sceptical of fact-checks that challenge things we thought were true and motivations to accept fact-checks (or any other information, true or false) that confirm them (Kahan 2017).

A defining feature of today’s informational environment is that such scepticism is not simply a matter of private reasoning. The challenge to fact-checkers and experts is immediate and public. Debates in TV studios and on social media expose citizens not only to ‘the facts’ but also to interpretations of those facts. Moreover, anyone can dispute an expert fact-check – and can expect positive reinforcement from partisan or ideological sympathisers, since the age of information abundance and fact-checking is also the era of polarised politics. Extremists and populists alike encourage mistrust of established institutions and their claims to provide objective fact-checking. Populist politicians have portrayed experts as self-interested actors, espousing views that fly in the face of the common sense or the personal experience of the people. In the much-quoted words of Michael Gove, a prominent Leave campaigner during the UK’s 2016 referendum on EU membership, “I think the people of this country have had

enough of experts with organisations with acronyms saying that they know what is best and getting it consistently wrong” (SkyNews 2016). In short, there is no lack of elite cueing of the views expressed in the tweets opening this article.

Yet we know surprisingly little about the persuasiveness of these views. Put another way, how much of the positive effect of fact-checks and corrections is undone when they are challenged in this way? And does the answer depend on the narrative undermining the facts? As illustrated by those sample tweets, post-truth rhetoric offers several such narratives, such as challenging the neutrality of the source, or asserting the primacy of personal experience or even personal feelings over evidence. Which are most persuasive? Existing research has not yet addressed these questions or, more generally, assessed the effectiveness of corrective information in the face of post-truth rhetorical challenge. Instead, studies have typically assessed the impact of fact-checking in ‘one-shot’ experiments. They have identified those who hold misperceptions, exposed them to a fact-check, and measured the accuracy of belief after the check Weeks (2015). This research design is appropriate for measuring the immediate impact of such a correction. However, by giving fact-checkers the final word, it does not replicate a world in which they come under immediate challenge and so cannot estimate the resistance of their corrections against such challenge.

It is not possible to capture the full cacophony of public debate in a controlled survey experiment. Nevertheless, we can go beyond existing studies by testing the robustness of fact-check effects against the immediate challenges observed in a post-truth environment. We do this with a survey experiment based on perceptions about immigration: an emotionally charged issue influential in recent UK general elections and the 2016 referendum on European Union membership (Sobolewska and Ford 2020). We first established people’s misperceptions about immigration and then treated them to corrective expert evidence countering whichever false claim they were most convinced was true. Then, before measuring the revised perceptions of our treatment groups, we exposed them to a shot of the kind of post-truth rhetoric that opened this article. These undermining messages did not directly refute the accuracy of the

fact-check but gave respondents one of three reasons to take the expert evidence ‘with a big pinch of salt’.

Our results support previous studies on the impact of fact-checks. Not everyone switched all the way but most shifted their perceptions in line with the fact check, becoming less convinced it was true. However, exposure to post-truth rhetoric then significantly undermined this effect. Compared to those receiving only the fact check, treatment group respondents moved almost one third of the way back towards their original misperceptions. This effect was consistently significant and varied surprisingly little in size, regardless of the strand of post-truth rhetoric or the credentials of the commentator.

This paper is structured as follows. First, we review the literature on resistance to factual correction and formulate four hypotheses designed to assess the impact of fact-checking and exposure to post-truth rhetoric. Then we describe the data and design used to test these hypotheses and present our results. Finally, we consider the external validity and implications of these findings, suggesting they may well understate the undermining impact of post-truth rhetoric given its likely repetition and amplification within people’s social networks.

Resistance to correction: drivers and limits

Humans tend to resist information that challenges their factual beliefs, even when it represents otherwise compelling testimony (Johnson and Seifert 1994; Wyer, Clore, and Isbell 1987). The drivers of this resistance are well-known. We are motivated reasoners, driven not only by accuracy but also by directional goals, powerful incentives to maintain existing opinions (Kunda 1990; Mercier and Sperber 2011; Bolsen and Palm 2019). This is partly because it is easier to stick to one’s opinions rather than engage in time-consuming consideration of incoming information. But it is not only about cognitive convenience. Many people are partisans, some are ideologues, and all belong to social groups, and these identities structure

and drive motivated reasoning (Nyhan and Reifler 2010; Bavel and Pereira 2018; Ecker and Ang 2019; Ecker et al. 2022). There is psychological discomfort in accepting information and forming beliefs in conflict with those held by fellow partisans or other group members – especially if those beliefs are central to the group’s identity or that group is central to our own social or political identity (Zaller 1992; Kahan 2017; Guess and Coppock 2018; Wittenberg and Berinsky 2020).

Within this framework, those fact-checking political claims can be seen not only as providing accurate information but also as priming accuracy goals. Their purpose is to encourage citizens to see that, amid the motivated to and fro of political argument, some claims are correct and some are not. The evidence on correction of misinformation suggests that it has a mixed record. There are prominent examples of success (e.g. Redlawsk, Civettini, and Emmerson (2010); Nyhan (2020); Porter and Wood (2019)) but also prominent examples of resistance to correction (Nyhan and Reifler (2010); Thorson (2016)). Reviews of this extensive literature report a similar mixed picture (see Lewandowsky et al. (2012); Flynn, Nyhan, and Reifler (2017); Swire and Ecker (2018); specifically on fact-checking, see Nieminen and Rapeli (2019)). There is now enough empirical work to support meta-analysis, and three recent analyses found that factual corrections significantly increased the accuracy of people’s perceptions (Chan et al. 2017; Walter and Murphy 2018; Walter et al. 2020). This effect is moderate in strength (a standardised effect size of around 0.3). It is much weaker in relation to ideologically divisive issues, and in election campaigns when partisan considerations are activated (Walter and Murphy 2018; Walter et al. 2020). In short, and not surprisingly, neither accuracy nor directional goals win the battle outright.

Put another way, fact-checks are a significant intervention in that battle but do not resolve it in favour of accuracy. They are not the last word. Many of those citizens momentarily persuaded by a fact-check retain at least some impulse to reject it, an impulse that can be activated by those with the incentive and platform to do so. The question is: how is it activated? What are the lines of counter-attack against factual correction, and how successful

are they? These issues have been much less explored, especially empirically. We know much more about why people might resist correction than about how.

The theoretical frameworks from previous work allow us to identify three potential mechanisms (Wittenberg and Berinsky 2020; Ecker et al. 2022). The first, labelled here ‘expert bias’, discredits the source of the fact-check. Numerous studies show that the perceived expertise, trustworthiness and general authoritativeness of a source strongly predicts the extent to which people accept their fact-checks and corrections (Guillory and Geraci 2013; Vraga and Bode 2017; Berinsky 2015; Swire et al. 2017). And the interventions of fact-checkers can be portrayed as just that – the facts: authoritative and objective contributions from those seeking only to inform the debate. However, a common counter-attack portrays them instead as partial in either or both senses of the term: at best, not telling the full story; at worst, biased contributions from those ideologically committed to one side of the debate (Ylä-Anttila 2018; McIntyre 2018). If sources can be thus discredited, their contributions can be dismissed.

The second way that checks can be undermined, which we label ‘invoking personal experience’, asserts the importance of personal experiences rather than objective evidence. Where statistics are at odds with personal experience or anecdotal evidence, people attach more weight to those things that they directly observe (Shelby and Ernst 2013). People are prone “to treat subjective experiences as truer than objective facts” (Kubin et al. 2021). Corrections can be dismissed as ‘your truth but not mine’ (see Bavel et al. (2021)).

That last phrase brings us to the third mechanism that may undermine fact-checks. This line, which we follow Stephen Colbert (2005) in labelling ‘truthiness’, suggests that intuition or gut feeling are as valid a route to the truth as statistics or scientific evidence. When seemingly contradictory statistics are already being bandied around in political debate, incoming facts can be dismissed as just more data (and the phrase ‘lies, damned lies and statistics’ is available to anyone seeking to dismiss inconvenient evidence). This can encourage citizens

to suppose that there is no single truth, just rival opinions with their own supporting data – and, in such an unstable climate, people may conclude that their feelings provide a surer basis than objective facts. This is in some ways the most troubling mechanism because it seems to be beyond reason. Nevertheless, it may account for standard defences that those who contest the facts ‘have a right to their opinion.’

These three mechanisms are illustrated in turn by the three tweets that opened this article. Those examples were not hard to find. All three mechanisms are readily picked out in the chorus of scepticism that typically greets fact-checking interventions on social media. There are frequent attacks on the neutrality and accuracy of the fact-checkers’ sources and suggestions that equally valid data exist pointing in the other direction (Brandtzaeg and Følstad (2017); Shin and Thorson (2017)). These lines are readily available in public discourse to those with reason to dislike the corrective information put in front of them. What is lacking so far – and provided in this article – is empirical evidence on their effectiveness.

There is related evidence that ‘objective’ statistical evidence can be swiftly undermined. The focus there has been on the effects of journalistic ‘false balance.’ Reporting the counterarguments of ‘maverick scientists’ reduces confidence in expert consensus and lends credibility to a minority view that is not supported by the evidence (Dearing 1995). Climate science provides vivid illustrations (McCright et al. 2016; Linden et al. 2017). In one study, respondents were presented with a pie chart stating that ‘97% of climate scientists have concluded that human-caused climate change is happening’. Some received an additional treatment: They were shown a (real) petition urging the U.S. government to reject the Kyoto protocol because ‘there is no convincing scientific evidence that human release of carbon dioxide... will... cause catastrophic heating of the Earth’s atmosphere.’ On its own, the pie chart had a positive effect on perceptions of expert consensus. When accompanied by the petition, it had no effect at all (Linden et al. 2017).

This raises the two major questions with which our experimental study extends research on

resistance to correction. If a fact-check is accompanied by a statement encouraging people to disregard the statistics, how far does this offset the positive effect of the fact-check? And does the answer depend on which of the three mechanisms – expert bias, invoking personal experience, or truthiness – is activated by that message? We also examine the potential importance of source credibility in these dynamics. Finally, we test whether these rhetorical challenges encourage more general ‘post-truth’ reasoning about the blurring of facts and opinions and the acceptability of ignoring ‘the facts.’

Hypotheses

We tested four pre-registered hypotheses ^{1 2}. The first, consistent with the weight of literature reviewed above, is that the expert factual correction will reduce misperceptions.

Fact-Check Hypothesis (H1): When respondents are asked to reassess a false statement after it is corrected by an expert fact-check, they are less likely to rate it as true.

Confirmation of H1 is a precondition for testing our second and most important hypothesis: that the positive effect of the fact check will be undermined by subsequent exposure to ‘post-truth’ rhetoric.

Undermining Message Hypothesis (H2): The effect of the fact-check is weaker among those who were exposed to an undermining statement.

We tested three post-truth rhetorical messages designed to undermine respondents’ reliance on expert evidence. These reflect the three mechanisms set out above (‘expert bias’, ‘invoking personal experience’ or ‘truthiness’). While each message is prominent in challenges to factual corrections, their relative persuasiveness has not been tested in past research and we

¹Registered with OSF at [link blinded for review].

²Note for reviewers: Unfortunately, our OSF submission includes a pre-analysis plan that is not anonymised. We have shared a link to our OSF pre-registration with the editors and include an anonymised version of the pre-analysis plan in the appendix.

had no a priori expectation that one argument would be more powerful than another. We simply anticipate that H2 will be confirmed for all three messages.

It seems only natural to suggest that the effectiveness of these rhetorical messages would vary according to how authoritative the individual that makes a statement is. We noted earlier that authoritative sources are more likely to correct misinformation; it seems reasonable to suppose that they can better discredit correct information, too. Testing this supposition and using credentials as a signal for authority, we hypothesise:

Authoritativeness Hypothesis (H2a). The effect in H2 is stronger if the undermining statement comes from a professor rather than a blogger.

Finally, we assessed the effect of exposure to undermining challenges on broader ‘post-truth’ thinking. Here, our experiment stretches well beyond existing research and so our hypotheses are more exploratory and worth setting up in more detail. The first point is that this may be a new contribution but it is on an age-old issue: citizens’ belief in objective truth and the value they attach to it. Writing about the Holocaust, Hannah Arendt suggested that the scale of lying by politicians had left the public at ‘the point where they would, at the same time, believe everything and nothing’. Those purveying mass propaganda ‘discovered that its audience . . . did not particularly object to being deceived because it held every statement to be a lie anyhow’ (Arendt 1951). Forty years later, Steve Tesich (1992) coined the term ‘post-truth’ in writing about the trauma of the Watergate and the Iran-Contra scandals: “We, by our actions, are saying that this is no longer necessary, that we have acquired a spiritual mechanism that can denude truth of any significance. In a very fundamental way we, as a free people, have freely decided that we want to live in some post-truth world”.

Yet there is a long distance between external observers describing the public as ‘post-truth’ and citizens themselves embracing the rejection of objective truth or endorsing the notion of ‘alternative facts’. There is little if any direct evidence of such post-truth reasoning. One important exception tests Brotherton and Son’s (2021) proposition that people might

metacognitively categorise uncongenial claims as ‘more opinion than factual’, and congenial claims as more factual than opinion. In a 2018 Pew survey, 5,000 U.S. Americans were presented with a battery of statements and asked to categorise each as a factual or opinion statement. A substantial proportion mislabelled factual statements as opinion statements – especially if those statements challenged their partisan identity. For instance, almost 40% of Republicans deemed the statement that ‘President Barack Obama was born in the United States’ as an opinion (Mitchell et al. 2018). Categorising a statement as opinion may also be a response to complexity: close to half of both Democrats and Republicans categorised ‘Spending on Social Security, Medicare, and Medicaid make up the largest portion of the U.S. federal budget’ as an opinion. This re-categorization of facts as opinion is one way to allow oneself to disregard the evidence, and thus also a means by which post-truth rhetoric can undermine that evidence.

Categorising factual corrections as matters of opinion is the first of three strands of post-truth reasoning that we hypothesise will be encouraged by the undermining rhetoric. The other two are more explicit statements of ‘truthiness’ inspired by a combination of Stephen Colbert’s (2005) introduction of the notion in 2005 and Kellyanne Conway’s notorious notion of ‘alternative facts’ (Todd 2017). The underlying conjecture is that, if encouraged to reject ‘the facts’ in this specific case, respondents might feel entitled or emboldened to do the same more generally.

‘Post-truth’ Reasoning Hypothesis (H3). Respondents who saw an undermining statement are more likely to:

- a. Categorise uncongenial factual questions as matters of opinion.
- b. State that, in this case, they ‘believe something different’ from the facts; and
- c. Agree in general that ‘it is okay to disagree with the facts’ if the facts point another

way.

Case, data and design

These hypotheses were tested in a survey experiment based on common misperceptions about immigration in Great Britain. This makes a useful case for several reasons. Immigration is subject to widespread misperceptions. This is especially true of perceptions about the numbers of immigrants, but it also extends to beliefs about the origins and status of immigrants (Duffy (2014); Hopkins, Sides, and Citrin (2019); Gorodzeisky and Semyonov (2016); Lutz and Bitschnau (2022)). The issue has also been thoroughly politicised in Britain (Sobolewska and Ford (2020)). Many citizens have powerful directional goals about immigration. These were closely tied to voting behaviour in the EU referendum (Hobolt, Leeper, and Tilley 2021). The referendum also linked immigration attitudes to powerful group identities. We might reasonably assume that an expert asserting a pro-immigration view might be labelled as a Remainer – and anyone challenging the same expert would be labelled a Leaver. There is ample evidence that beliefs are harder to shift if they are seen as central to an individual’s group identity (Kahan (2017); Nyhan (2020)). Messages that encourage the rejection of unsolicited claims about immigration should therefore fall on receptive ears.

The main survey was fielded online in June 2019 on a sample of 2,936 British citizens. Fieldwork was by Deltapoll, an internet survey agency that recruits from online panels. This was a non-probability sample but is representative of the British electorate on key demographic and political variables. (Sample characteristics are reported in Appendix A.)

The main survey was preceded by a pilot study with a diverse convenience sample of 200 British adults recruited via Prolific. We pre-tested various false claims on both sides of the immigration debate and used the results to choose two false claims widely believed on each side. The ‘pro-immigration’ claims were: ‘Immigration to the UK does not affect the wages of the low-paid,’ and ‘The majority of crimes in London are committed by white people,

not ethnic minorities.’ The ‘anti-immigration’ claims were: ‘There has been a sharp rise in the number of people applying for asylum in the UK in the past ten years,’ and ‘European immigrants receive more in benefits and services than they pay in taxes.’

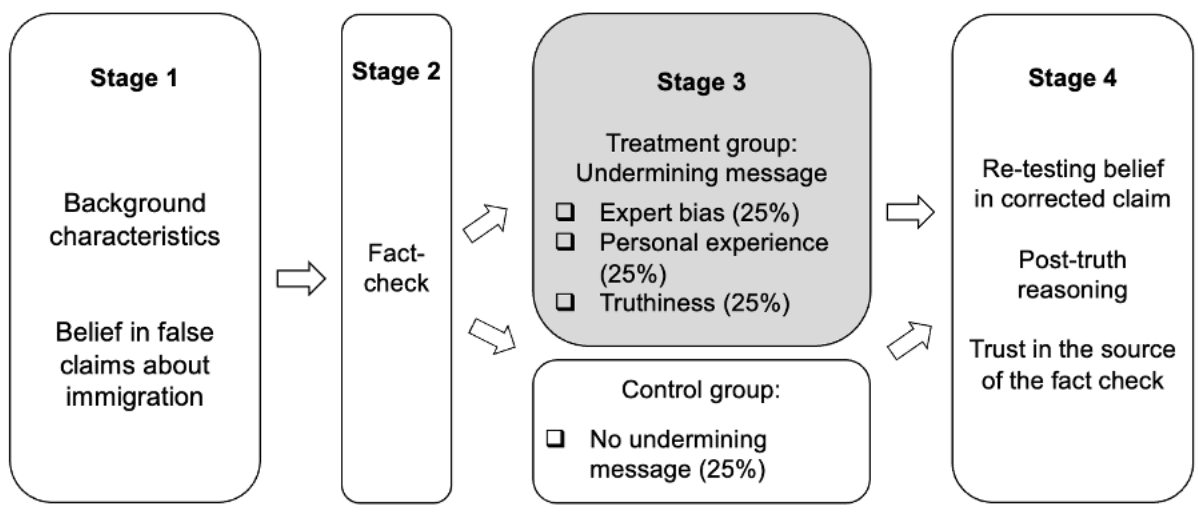


Figure 1: Design of the main experiment

The survey opened with background questions, including socio-demographics and views about immigration. The experimental phase had four stages, as shown in Figure 1. In **Stage 1**, we tested belief in the false claims. Respondents were presented with eight statements: four about immigration and, to disguise our intent somewhat, four distractor items about issues like fracking and plastic bag usage. Two statements were common misperceptions among immigration critics (“There has been a sharp rise in the number of people applying for asylum in the UK in the past ten years., and”European immigrants receive more in benefits and services than they pay in taxes.”). The other two were common misperceptions among pro-immigration respondents (“The majority of crimes in London are committed by white people, not ethnic minorities.”; “Immigration to the UK does not affect the wages of the low-paid.”) Respondents rated each claim on a seven-point scale from ‘definitely true’ to ‘definitely false.’ We used the true-false labelling (rather than, say, an agree-disagree scale) to encourage respondents to see these as matters of fact rather than opinion, and a scale rather than a binary (‘true’ v. ‘false’) question to allow respondents to indicate not just

their assessment of those statements but also their confidence in their own assessment. The repetition captures how far that confidence is affected by a fact-check and its subsequent undermining.

Stage 2 was a fact-check contradicting the false immigration claim that they had rated closest to ‘Definitely true’. If two or more items were tied, we randomly assigned respondents to one of the relevant fact-checks. That means that the specific information received at this stage depended on the answers at Stage 1. We attributed all four fact-checks to the same (fictional) person. The introduction read, ‘We have asked Richard Clarke, Professor of Economics at the University of Oxford to provide us with information about the statements you just read.’ We chose a generic British name and a photo of an austere, nondescript, middle-aged man in front of a blackboard. We gave him impressive credentials to boost his authority: a PhD from Harvard University and consultancy for the Office of National Statistics, a body widely trusted in Britain (Morgan 2019). The content of the fact-checks was based on real statistics and mimicked the general approach of fact-checking sites such as Full Fact.

At **Stage 3**, the sample was split. Treatment-group respondents (75%) were presented with a message undermining the factual correction while the control group (25%) skipped directly to Stage 4. The undermining message was introduced as ‘a comment on that information from a different source.’ It was attributed to ‘David Williams’, an avuncular, grey-haired man wearing a beret and holding a cigar. (To avoid confounding effects of age, ethnicity, or gender we chose pictures of middle-aged white men for both the expert and the commentator.) Williams was described either as a ‘Retired Professor at the London School of Economics’ (high authority) or as a ‘Blogger’ (low authority).

In a 3 x 2 design among treatment-group respondents, we varied both the content of the message and the authoritativeness of the source. The statement always started with the words, ‘I would take these statistics with a big pinch of salt.’ The variations that followed

were designed to capture the three mechanisms described earlier. While they are more formal in tone than the example tweets quoted at the outset, they deploy very similar undermining narratives:

Expert bias: ‘The fact that someone is a professor doesn’t mean that they don’t have an agenda. And we all know that there is a lot of scope to choose and present statistics so that they end up saying just what you want them to say.’

Invoking personal experience: ‘A graph might say one thing but the experience of people’s everyday lives could be quite different. And I think that a lot of people reading those statistics will say: that doesn’t sound like the world I live in.’

Truthiness: ‘There’s so much information and so many statistics out there that it can be hard to know what to believe. In that case, I think it’s best to trust your instincts even if it looks as if the facts are different.’

At **Stage 4** all respondents were asked to rate the corrected claim again on the seven-point true-to-false scale. The difference between these ratings and the pre-correction ratings is the main quantity of interest in our analysis. This stage also included three questions examining why the more resistant respondents blocked out the information from the fact-check and gauging whether these strategies were encouraged by the undermining challenges. Here, we were particularly interested in whether there were signs of reasoning that might be deemed ‘post-truth’: that is, evidence that people who were wrangling over the facts would question whether there was such a thing as ‘a fact’. First, we asked respondents to rate each of the two (false) statements on a 0-6 scale from ‘Purely a matter of fact’ to ‘Purely a matter of opinion’. Next, we asked whether respondents thought that the statistics that the fact-checker had provided were consistent with what they had believed beforehand. Those answering ‘No’ were then asked to choose which of three statements best described their position: a) ‘The statistics are probably right but I believe something different’; b) ‘I think that the statistics are wrong’; or c) ‘The statistics made me change my mind’. Then, in perhaps the most direct

approach to post-truth thinking, we asked whether respondents agreed or disagreed with the general proposition that ‘It is OK to disagree with the facts if that’s what you believe.’ All these questions were asked of both treatment and control group respondents.

Finally, we asked two questions about the source of the factual correction (the Oxford professor) and, where applicable, parallel questions about the source of the undermining message (the LSE professor or the blogger): how much respondents trusted the sources, and how much they trusted their information.

Participation in this study took just over 5 minutes on average; respondents were remunerated according to Deltapoll’s standard practice. Participation was strictly voluntary; respondents were asked for consent and could withdraw at any time. The fact-checks were factually accurate; the only deception we used was to attribute the fact-checks, as well as the undermining messages to fictional people.

Results

Before reporting our results, we should note two points about our approach and its implications for those results. The first concerns asymmetries in the sample in terms of overall preferences about immigration and in terms of prior belief in the false claims. In line with other surveys, there was variation in overall preferences but a discernible skew towards preferences for a reduction in immigration to the UK (Schwartz et al. (2021)). About a third (32%) were content with current levels, about half (47%) preferred fewer, and about a fifth (21%) preferred more immigrants. Unsurprisingly, the false anti-immigration claims were more widely believed and, therefore, more often provided the basis for fact-checking. They were also more strongly believed. Almost half of those who were directed to an expert challenging one of the two anti-immigration claims had initially rated the respective claim as ‘definitely true’ (44% of those who learned about the number of asylum seekers, 49% of those who learned about the cost of immigration). This was true of just 30% of those who were

directed to an expert challenging one of the two pro-immigration claim (26% of those who learned about crime in London and 34% who read about low-paid wages).

The second, related point is a cautionary note about the motivations that respondents brought to the treatments. The statements that were corrected were chosen based on specific beliefs rather than underlying ideology or motivations. Our design assumes that those who were most sceptical of immigration would believe most strongly in one of the false anti-immigration claims, and the reverse would be true of pro-immigration respondents. This was mostly but not always the case. Overall, just under half of our pooled sample (47%) were directed to a fact-check that challenged a statement on their side of the immigration debate, a fifth (21%) were directed to one that challenged a statement on the other side of the immigration debate, and a third (32%) took no clear side, neither wanting to increase or decrease immigration. Importantly for our purposes, these numbers differed considerably across the four fact-checks: Of those directed to one of the two anti-immigration fact-checks, 62% had expressed anti-immigration views (25% neutral, 14% pro). In contrast, of those directed to one of the two pro-immigration fact-checks, only 33% had expressed pro-immigration opinions (45% neutral, 22% anti). Given the difficulty of accepting evidence that challenges one's partisan or group identity this means that our anti-immigration sample had a stronger motivation to reject the evidence (Lodge and Taber 2013; Kahan 2017; Nyhan and Reifler 2019).

Of course, preferences on immigration numbers are not a perfect measure of underlying attitudes and not a robust basis on which to discard part of the sample at the outset. It is safer simply to bear in mind that our full-sample analyses make for a conservative test of the effect of that undermining challenge, because of this hefty chunk of the sample that had a weaker incentive to respond to it.

Effects of the fact-check and the undermining message

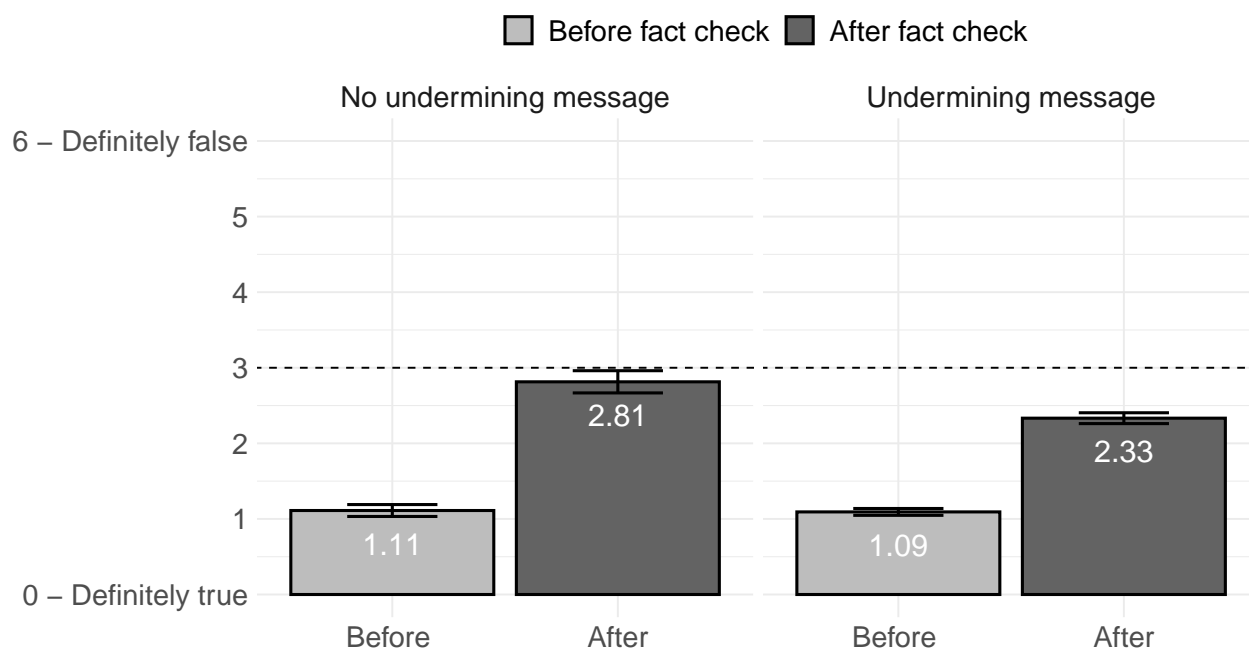


Figure 2: Average assessment of the false claims by exposure to an undermining message. Mean and 95% confidence intervals.

As anticipated in H1, our fact-checks did indeed significantly reduce belief in false claims. Figure 9 in the appendix shows that all four corrections had roughly similar effects, which warrants our reporting analyses pooling across the four in the remainder of this article. Those pooled data for the control group are shown at the left-hand side of Figure 2 (‘No undermining message’), which shows mean assessments of the respondents’ key claim (on the 0-6 scale from ‘Definitely true’ to ‘Definitely false’) before and after the authoritative correction. On average, the fact-checks shifted respondents about 1.7 points closer to the ‘false’ end of the scale, from 1.11 (i.e., barely a point above 0 – definitely true) to 2.81 (i.e., almost reaching the mid-point of the scale, as denoted by the dotted line.) This result – the experts’ fact-checks having considerable impact but leaving plenty of people unconvinced – tallies with the findings of partial success reported earlier from past research. Indeed, given that this was a one-shot challenge of factual beliefs about a politicised issue that engaged strong attitudes, a shift of 1.7 points on a seven-point scale is on the large side relative to

previous studies (Walter et al. 2020).

Our key questions are: was that success undermined by post-truth rhetoric and, if so, how far? Comparing the control group and treatment group responses reveals clear support for H2. Exposure to an undermining message did indeed offset a significant chunk of the positive effect of the fact-check. Among treatment group respondents recommended to ‘take those statistics with a pinch of salt’, the fact check did have an effect but moved respondents only 1.24 along the scale (from 1.09 to 2.33, as shown at the right-hand side of Figure 2). If this 1.24 is compared to the 1.70-point effect among control group respondents, we can say that the undermining message undid almost a third ($1-1.24/1.70$, or 27%) of the fact-check effect. The difference between these mean shifts is statistically significant ($t(1088)=5.53$, $p<0.01$) and moderate in size by conventional yardsticks, amounting to around a quarter of a standard deviation (Cohen’s $d=0.26$). Effect size is partly a matter of interpretation and expectations, however, and in the concluding discussion we assess whether this is a strikingly high or reassuringly low impact of the undermining message.

These mean differences conceal instructive information about the impact of the fact checks and the subsequent undermining messages. In either case, it could be that a few people’s minds are changed completely while others remain unmoved, or that most people are nudged along the scale by the evidence and arguments. Comparing the top two panels of Figure 3 shows that fact-checks had a particularly pronounced effect on the share of respondents who thought the false statements were ‘definitely true’: this halved from 40 per cent (before the fact-check) to 19 per cent (after the fact-check) among the control group who offer our pure test of the effect of the fact-check. Around one in six respondents (17%) were so convinced by the expert fact-check that they re-rated the false claim as ‘6 - definitely false’. If we then compare the bottom two panels (both after the fact-check), we can see the effect of exposure to a message from another professor, or a blogger saying they would ‘take those statistics with a pinch of salt’. The most notable difference is that proportion of the convinced, which was just 7% among those who had seen one of these undermining messages. It is also noteworthy

that that the treatment did not send respondents all the way back to ‘Definitely true’. The effect of post-truth rhetoric was not to instil conviction but rather to encourage doubt and tilt respondents back towards their original belief.

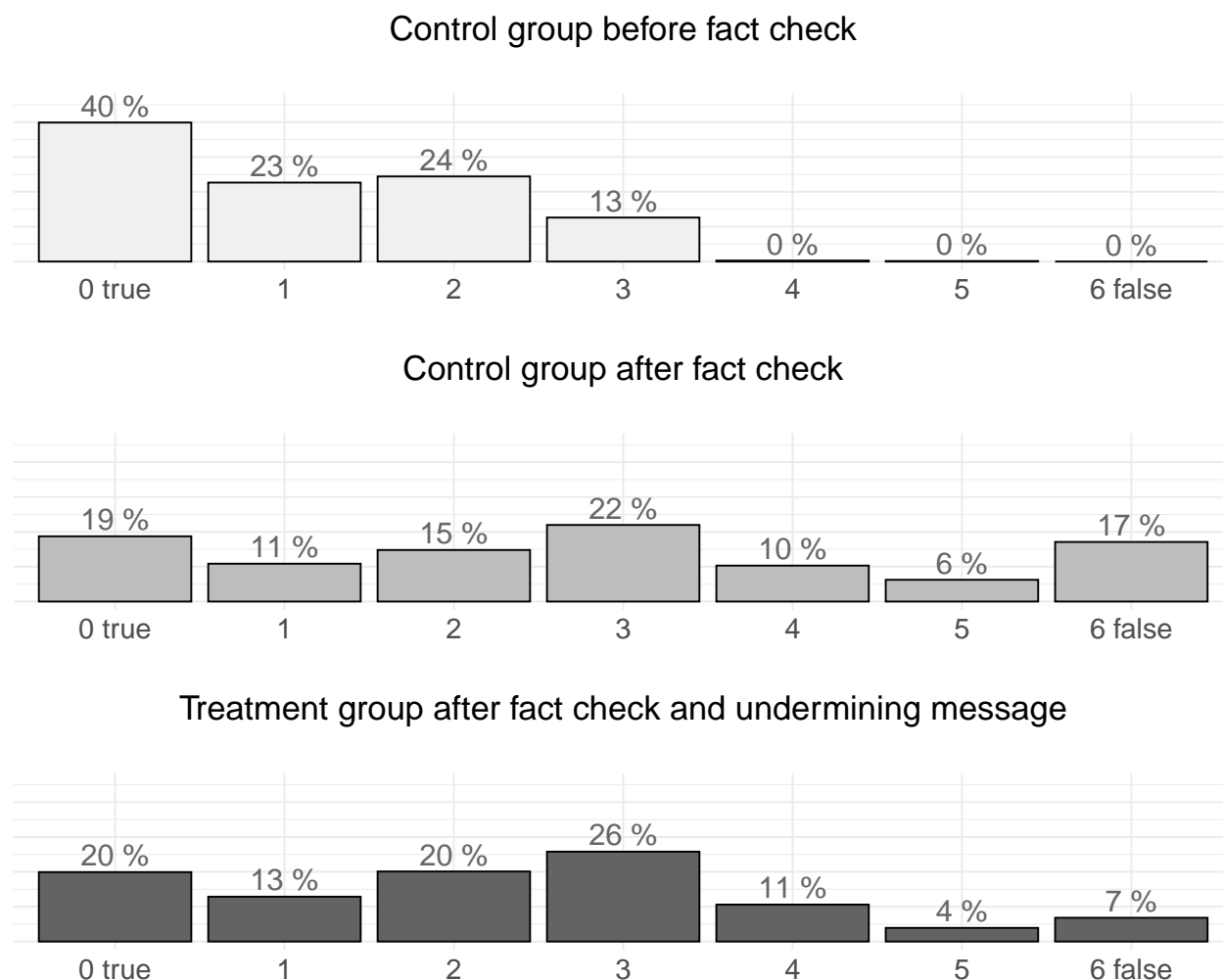


Figure 3: Assessment of the false facts along the full 0-6 scale: pre-treatment (all respondents); post-treatment (control group); post-treatment (treatment group).

One of the points we raised at the outset of the results section concerned the tendency for anti-immigration beliefs to be more strongly held. Consistent with that, those who had seen an anti-immigration claim corrected in a fact-check were more sympathetic to

the undermining challenge. Among this group, 31% of the effect of the fact-check was undone. This proportion was down to 19% among the smaller group who had seen a pro-immigration claim corrected. This gap largely reflects differences in the strength of prior perceptions. As shown in Figure 9 in the appendix, those seeing one of the anti-immigration claims corrected started out with somewhat stronger misperceptions on average, hovering around 1 on the 0 (true) - 6 (false) scale (0.9 and 1.1, respectively), while those seeing one of the two pro-immigration claims corrected started with an average of 1.2 and 1.5, respectively. That means that our anti-immigration sample had more incentive to follow the lead of the person encouraging them to ‘take those statistics with a pinch of salt’. There is a subsidiary point about political context here. An anti-immigration opinion corrected by an academic statistician more closely reflects the nature of an EU referendum campaign in which, as illustrated by the earlier quote from Michael Gove, the notion of ‘expertise’ became discursively associated with the immigration-friendly Remain side and hence more often questioned by immigration-sceptic Leave campaigners and voters. The content of the undermining challenge, pointing to biased academics, noting that the official information did not match their personal experience, or encouraging people to trust their instincts more than statistics is probably more familiar and more popular with those who would like to reduce immigration.

The second point is to reiterate that these figures represent a conservative estimate of the effect of these undermining messages because of that one fifth of the sample whose most confident false belief ran against the grain of their immigration preferences. Those respondents had more incentive to embrace the fact-check information as a congenial surprise, and less incentive to respond to the undermining message. If they (and a few more who did not see immigration as an important issue) are removed from the sample, the countering effect of that undermining message among those remaining (n=2054) rises to 29% overall (31 and 32% in the anti-immigration sample; 19 and 29% in the pro-immigration sample, as shown in Figure 10 in the Appendix).

Interestingly, however, this same sub-sample whose immigration preferences gave them a reason to reject the facts and accept the challenge was *not* more likely to reject the initial fact-check. They moved 1.67 points closer to ‘false’ after seeing the statistics (from 0.91 to 2.58) – only marginally below (and not statistically significantly different from) the 1.7-point shift in the full sample. In other words, the more motivated were just as ready to listen to the expert when he had the final word but, when he was challenged, they were particularly inclined to return to their initial belief. This is an important insight both for psychologists studying motivated reasoning and for fact-checkers.

Effect of an undermining message by source and content

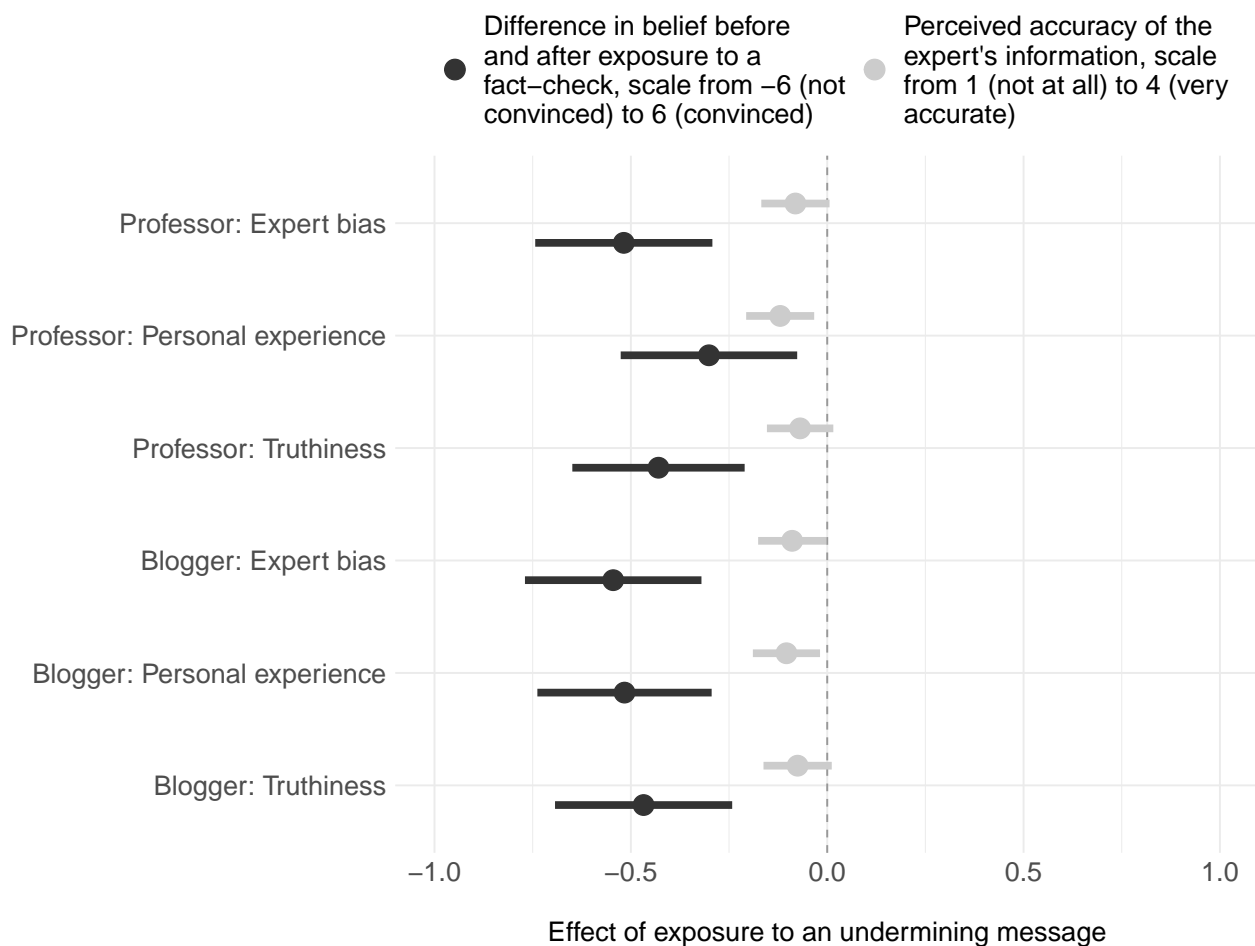


Figure 4: Effect of the undermining message on the effect of the fact-check.

The next step is to compare the arguments that convinced some respondents to dismiss or at least to doubt the fact-check. Were all three effective? Were some more effective than others? The respective answers, based on Figure 4, are ‘yes’ and ‘not really’. The figure shows coefficients from OLS regression models estimating the effect of exposure to any of the six versions of the undermining message. (Full model results are reported in the appendix.) The first estimate in each pair (printed in black) is the key measure testing H2: how far does the undermining message shift ratings of the original belief on the true-false scale (relative to the control group who see only the fact check)? Scores are negative here because those messages move respondents down the scale towards 0 (‘definitely true’) and centred around -0.46 which was the average difference among the whole sample. The scores actually vary rather little around that -0.46. All the messages tested undermined a statistically significant chunk of the effect of the fact-check but there is no significant difference by message content. There is some sign that invoking personal experience is least effective and claiming expert bias is most effective, but the differences are small. Whether comparing pairs of the means in Figure 4 or collapsing across sources for an overall ANOVA test of the effect of challenge message, there is no route to a statistically significant result.

There is another null finding, contrary to H2a, when it comes to the source of the non-factual challenge. Only on one of the three messages (‘Personal Experience’) was there even a sign of the professor carrying more weight. It may be that the authoritativeness of the source matters less when the content is more ‘everyman’: no expertise is required to tell people to trust their experience and instincts (and not much is required to know that professors may be biased!). In any case, an uncharitable summary is that respondents clutched at any straw they were offered, by whomever, to discredit statistics challenging their views.

The second estimate in each pair (printed in grey) bears on one potential mechanism through which these challenges could work: by undermining trust in the fact-checker and his evidence. Here, we present evidence from the question asking about the perceived accuracy of the expert’s information. The pattern is similar (see table 3 in the Supporting Information) if

we look instead at more general trust in that expert. That pattern is that, while post-truth rhetoric did leave a small dent in trust in the fact-checker and his information, the effect was relatively small and not always statistically significant at conventional levels. This is consistent with another point which is that, even in the treatment groups, trust in the fact-checker and his information remained on the trusting side of the scale – and significantly higher than trust in the post-truth commentator and their information.

Effect of an undermining message on post-truth reasoning

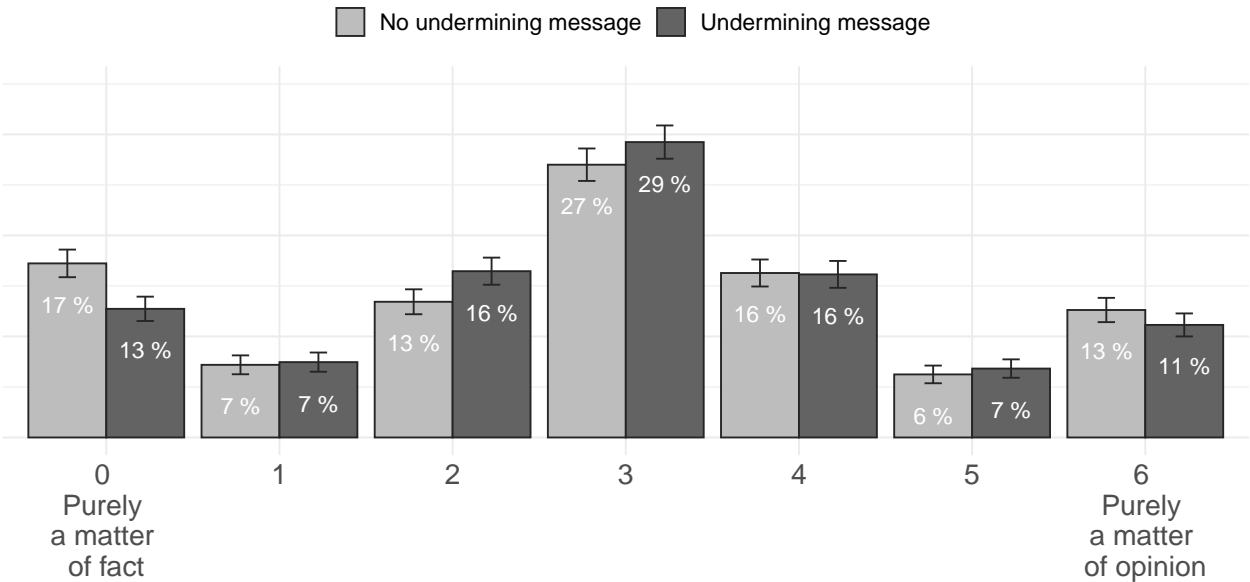


Figure 5: Assessment of the corrected statements on a fact-opinion scale by exposure to an undermining message.

How can we reconcile respondents’ drift back to their original belief after exposure to the undermining message and their reluctance to dismiss the fact-check and its source? One possibility is that the effect of post-truth rhetoric is to encourage people to believe that both that original belief and the (intended to be) contradictory fact check can be correct – or, at least, are both reasonable points of view. This brings us neatly to H3 and the notion

that the treatments encouraged a post-truth mindset. The first test is via that scale asking respondents to locate their key claim on a scale from ‘purely a matter of fact’ to ‘purely a matter of opinion’. We asked this question at the end of the survey. Figure 5 shows that there is very little sign of the undermining narratives nudging respondents up the scale to the latter view. Combining points 4, 5 and 6 on the scale, we can see that about one third reported that the statements were on balance a matter of opinion – but then this was also true among control group respondents in whose case the expert had had the final word. Clearly, as in the examples cited earlier from the 2018 Pew survey many citizens habitually regard political claims as opinions (Mitchell et al. 2018).

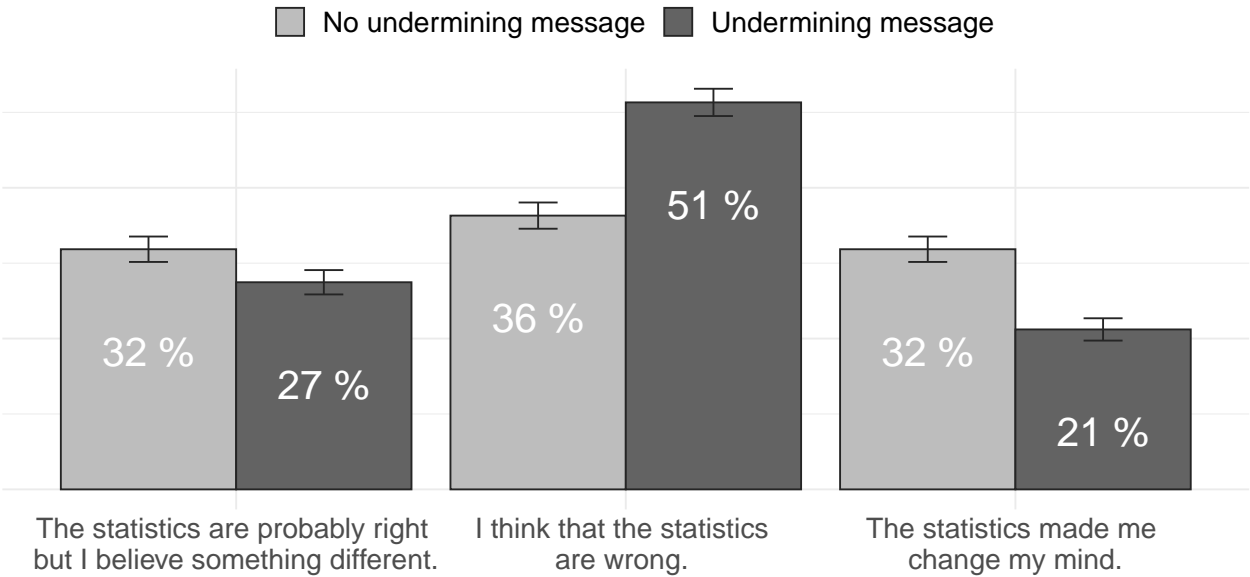


Figure 6: Responses to the question of where respondents stood by exposure to an undermining message.’

Next, we asked respondents whether the statistics they had seen were consistent with what they had believed. Despite the fact that the fact check statistics were explicitly presented as contradicting respondents’ initial beliefs, only 46% said ‘No’. That subset of respondents was then presented with three statements about this contradiction and asked to pick the one that best described where they stood. Again, the control group results are striking. Even without any post-truth nudge, the most common response (36%) was simply to assert that the expert’s statistics were wrong (see Figure 6). Less than a third (32%) said that the

statistics had led them to change their mind. Exposure to an undermining message did have a significant impact on responses ($\chi^2(2)=24.4$, $p<.001$), increasing the share of respondents who dismissed the statistics as ‘wrong’ and reducing the share of respondents who reported changing their mind. Interestingly, however, those messages did not lead respondents to endorse the post-truth compromise position that, while the statistics were ‘probably right’, they believed ‘something different’. There was some support for this position (despite its aroma of ‘alternative facts’) but it was not encouraged by undermining messages, whose effect was instead to trigger rejection of the statistical evidence.

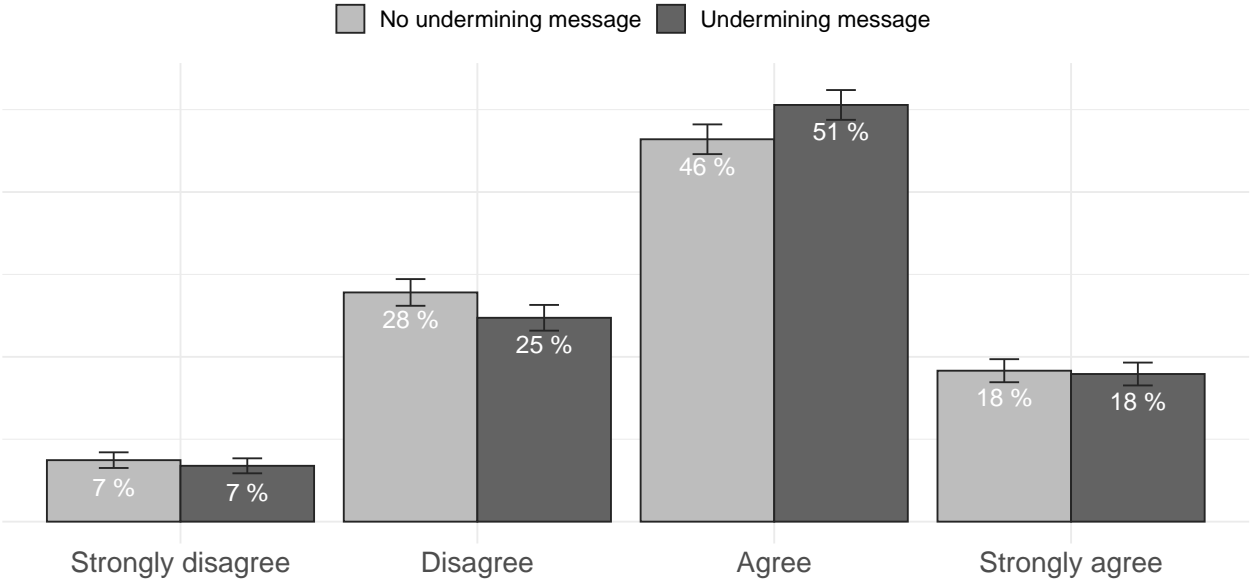


Figure 7: Agreement that ‘It’s OK to disagree with the facts if that’s what you believe’ by exposure to an undermining message.

Figure 7 shows reactions to the final indicator, the statement that “It’s OK to disagree with the facts if that’s what you believe.” There is a hint that the undermining challenge nudged respondents into agreement with that statement but the difference is just a few percentage points and the chi-squared test for the overall comparison falls well short of significance ($\chi^2(3)=4.4$, $p=.22$). One reason may be that almost two thirds of our control group sample (64 per cent) agreed or strongly agreed with the statement. What we had envisaged as a rather controversial claim, that respondents would need to be talked into by the undermining challenge, turned out to be a commonplace view. Throughout this section, the degree to

which even the control group agreed with ‘post-truth’-flavoured statements far exceeded our expectations – and rather limits the scope for treatment group respondents to agree even more strongly.

Finally, we test whether these largely null findings conceal at least some differences across the narratives in their capacity to prompt post-truth reasoning. Figure 8 shows the effect of exposure to each of the three undermining messages on the three key post-truth outcomes analysed above. In the case of the two scales (see Figures 5 and 7), the plot shows the OLS coefficient estimating the effect of the treatment on the mean response. With the categorical outcome variable (Figure 6), we estimate a logistic model and plot the treatment effect on the predicted probability of choosing the post-truth flavoured option ‘The statistics are probably right but I believe something different’. For the most part, the nulls persist, and there is little sign of differences across the three treatment narratives. The only exception is that exposure to the ‘Truthiness’ message had a significant but, contrary to our expectations, negative effect on the probability of choosing that ‘but I believe something different’ option. However, this is really just a restatement of this treatment’s especially positive effect on the probability of responding that ‘the statistics are wrong’ (Figure 6), a rather less post-truth-flavoured conclusion.

Concluding discussion

From one perspective, the results presented here are reassuring. Evidence from experts is usually heard, even if counter-framed by those seeking to undermine it. Fact-checks made some respondents abandon and many others at least feel less confident in their misperceptions. This was true even of a case like immigration, where many people had strong

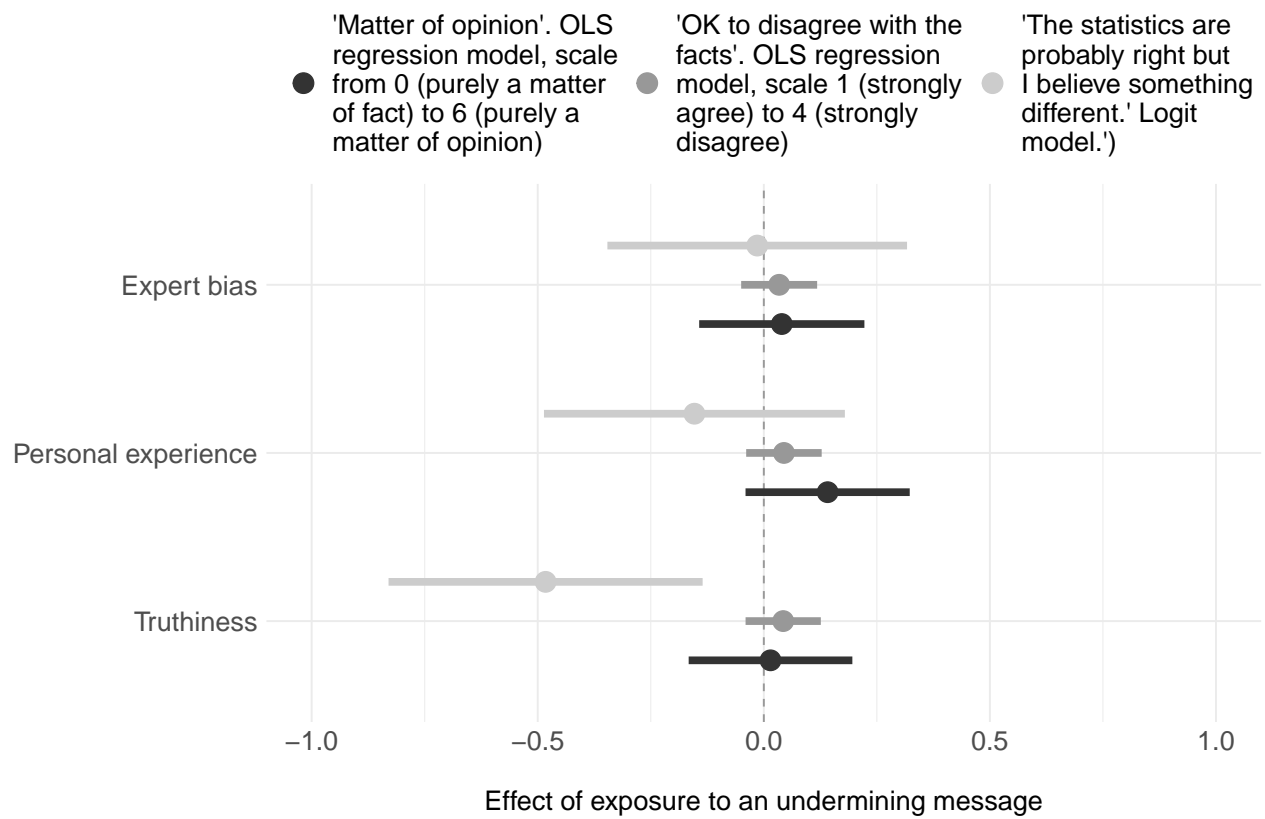


Figure 8: Effect of the undermining message on post-truth reasoning.

motivations – rooted in ideology and identity – to reject the factual correction. Given the importance of priors in information processing, we would not expect one piece of information to catapult people from one pole of the scale to the other. If a fact-check instead persuades people that they may not be as well informed as they thought, and perhaps induces them to seek out more information, then its work is well done.

Even where respondents were not just internally motivated but also externally encouraged to reject the fact-check, much of the work of fact-checking remained intact. The percentage of that work undone by a non-factual challenge varied somewhat by subsample but was never more than around one third. Even given a face-saving and cognitively easy excuse to ignore the fact-check, many did not do so. This is in stark contrast to one study we cited when the counter-information eliminated all positive effects of the factual treatment (Linden et al. 2017). One possible reason for the difference is that their dependent variable – perceptions of elite consensus on climate change – mattered less to respondents and so was more malleable. Another is that their countering treatment invoked scientific claims in a way that ours – intentionally – did not.

But there is a less optimistic reading of our results. It begins by emphasising that we gave more weapons to one side in our experiment. The undermining challenges were much briefer than the fact-checks and contained no graphical or supporting information. They were just a one-shot intervention from a lightly-credentialled individual. This is probably a bigger threat to external validity than the similarly one-shot nature of the fact-checks. In the real world, factual corrections – whether from fact-checkers, news articles or other sources – are often the single interventions then greeted with a chorus of disapproval along the lines of our non-factual challenges and the illustrative tweets that opened the article. Such ‘pile-ons’ are a reminder that public opinion is the product of social forces (Zaller 1992). The undermining chorus will often involve some combination of like-minded politicians, commentators, friends, and family – more trusted than the unknown challenger in our experiment. Our fact-checker was in that sense a more realistic representative of his real-world equivalent than was his

subsequent challenger. If one such undermining intervention can overcome nearly one third of a fact-check’s effect, perhaps a flurry of interventions could eliminate it entirely.

When the scenario is extended like that, another of our findings – that all three of the undermining messages had a significant effect – becomes more pertinent. The flurry of challenges to a fact-check is likely to include attacks from all sides, including the three mechanisms that we tested here. They may appeal to different audiences but, judging by these results, they all have an audience. However, it is hard here to adjudicate between two possibilities: the content of the message matters and all three worked; or the content of the message matters less than its presence. There is indirect support from the latter in the null finding on the source of the undermining message. It seems that some respondents needed very little encouragement to disregard unwelcome information. It just needed someone – anyone – to suggest that the statistics should be rejected. That said, we could have tested this ‘source irrelevance’ hypothesis rather further because, while post-experimental analysis did (mercifully) show appreciable gaps in trust and accuracy between the blogger and the professor, the former’s ratings were still quite high in absolute terms. Future research ought to test the effect of undermining messages by individuals who are more clearly non-authoritative – perhaps someone who openly claims that they have never investigated the statistics.

Among the limitations of our study, one prominent point is the difficulty of interpreting those responses that we suggested had a flavour of ‘post-truth’ reasoning. Prima facie, the fact that two thirds of British adults agree that “it is OK to disagree with the facts if that’s what you believe” looks like a ringing endorsement of Yuval Noah Harari’s (2018) claim that humans are a post-truth species. Yet they must be taken in context. Respondents had been shown a factual correction of a statement that they had expressed strong belief in and quite possibly held dear. This provided an unusually strong incentive to agree with that statement. (Had the fact-check been designed to confirm respondents’ beliefs, they may have answered the follow-up question about disagreeing with the facts very differently.) It also raises a question about what respondents read into ‘disagree with the facts.’ A recurring feature

of misperceptions research is that, even when corrections change beliefs, they are unlikely to change underlying attitudes (e.g., Hopkins, Sides, and Citrin (2019)). If respondents interpreted this statement as “it is OK to maintain your overall opposition to (support for) immigration even if this particular negative (positive) belief about immigration was wrong”, then there is much less of a post-truth flavour to their responses.

Even with these caveats, though, the proportions agreeing with those statements – presumably in the face of social desirability pressures pushing the other way – are striking enough to justify a research programme that defines and measures the tendency to post-truth reasoning. This is related to but different from the standard ‘belief in science’ measures. To provide a more nuanced understanding of general acceptance of post-truth arguments, future research ought to investigate post-truth thinking in different contexts, following welcome information, unwelcome information, or no information at all. Given the likely importance of social and ideological networks in vindicating such thinking, future experimental work could usefully gauge the impact of undermining messages such as ours when made by people with strong co-identities and shared group memberships.

Acknowledgements

This work was supported by the British Academy [SRG\171311]. The data and analysis underpinning this article will be available via [link blinded for review]. The authors declare no competing interests.

References

- Arendt, Hannah. 1951. *The Origins of Totalitarianism*. Schocken Books.
- Bavel, Jay J. Van, and Andrea Pereira. 2018. “The Partisan Brain: An Identity-Based Model of Political Belief.” *Trends in Cognitive Sciences* 22. <https://doi.org/10.1016/j.tics.2018.01.004>.
- Bavel, Jay J. Van, Diego A. Reinero, Victoria Spring, Elizabeth A. Harris, and Annie Duke. 2021. “Speaking My Truth: Why Personal Experiences Can Bridge Divides but Mislead.” *Proceedings of the National Academy of Sciences of the United States of America* 118: 1–3. <https://doi.org/10.1073/pnas.2100280118>.
- Berinsky, Adam J. 2015. “Rumors and Health Care Reform: Experiments in Political Misinformation.” *British Journal of Political Science*, 1–22. <https://doi.org/10.1017/S0007123415000186>.
- Bolsen, Toby, and Risa Palm. 2019. “Motivated Reasoning and Political Decision Making.” *Oxford Research Encyclopedia of Politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.923>.
- Brandtzaeg, Petter Bae, and Asbjørn Følstad. 2017. “Trust and Distrust in Online Fact-Checking Services.” *Communications of the ACM* 60. <https://doi.org/10.1145/3122803>.
- Brotherton, Robert, and Lisa K. Son. 2021. “Metacognitive Labeling of Contentious Claims: Facts, Opinions, and Conspiracy Theories.” *Frontiers in Psychology* 12 (March). <https://doi.org/10.3389/fpsyg.2021.644657>.
- Chan, Man-Pui Sally, Christopher R. Jones, Kathleen Hall Jamieson, and Dolores Albaracín. 2017. “Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation.” *Psychological Science* 28: 1531–46. <https://doi.org/10.1177/0956797617714579>.
- Colbert. 2005. “The Colbert Report.” *Comedy Central US*. <https://www.cc.com/video/61f6xj/the-colbert-report-intro-10-17-05>.
- Dearing, James W. 1995. “Newspaper Coverage of Maverick Science: Creating Controversy

- Through Balancing.” *Public Understanding of Science* 4: 341–61. <https://doi.org/10.1088/0963-6625/4/4/002>.
- Duffy, Bobby. 2014. “Perceptions and Reality: Ten Things We Should Know about Attitudes to Immigration in the UK.” *LSE Politics and Policy*. <https://blogs.lse.ac.uk/politicsandpolicy/perceptions-and-reality-ten-things-we-should-know-about-attitudes-to-immigration-in-the-uk/>.
- Dustmann, Christian, Tommaso Frattini, and Ian P. Preston. 2013. “The Effect of Immigration Along the Distribution of Wages.” *Review of Economic Studies* 80 (January): 145–73. <https://doi.org/10.1093/restud/rds019>.
- Ecker, Ullrich K. H., and Li Chang Ang. 2019. “Political Attitudes and the Processing of Misinformation Corrections.” *Political Psychology* 40: 241–60. <https://doi.org/10.1111/pops.12494>.
- Ecker, Ullrich K. H., Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K. Vraga, and Michelle A. Amazeen. 2022. “The Psychological Drivers of Misinformation Belief and Its Resistance to Correction.” *Nature Reviews Psychology* 1. <https://doi.org/10.1038/s44159-021-00006-y>.
- Flynn, D. J., Brendan Nyhan, and Jason Reifler. 2017. “The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs about Politics.” *Political Psychology* 38: 127–50. <https://doi.org/10.1111/pops.12394>.
- Fukuyama, Francis. 2017. “The Emergence of a Post-Fact World.” *Project Syndicate*.
- Gorodzeisky, Anastasia, and Moshe Semyonov. 2016. “Not Only Competitive Threat but Also Racial Prejudice: Sources of Anti-Immigrant Attitudes in European Societies.” *International Journal of Public Opinion Research* 28. <https://doi.org/10.1093/ijpor/edv024>.
- Graves, Lucas, and Federica Cherubini. 2016. “The Rise of Fact-Checking Sites in Europe.” Reuters Institute.
- Guess, Andrew, and Alexander Coppock. 2018. “Does Counter-Attitudinal Information

- Cause Backlash? Results from Three Large Survey Experiments.” *British Journal of Political Science*, 1497–1515. <https://doi.org/10.1017/S0007123418000327>.
- Guillory, Jimmeka J., and Lisa Geraci. 2013. “Correcting Erroneous Inferences in Memory: The Role of Source Credibility.” *Journal of Applied Research in Memory and Cognition* 2: 201–9. <https://doi.org/10.1016/j.jarmac.2013.10.001>.
- Harari, Yuval Noah. 2018. *21 Lessons for the 21st Century - Yuval Noah Harari*. Jonathan Cape. Random House.
- Hobolt, Sara B., Thomas J. Leeper, and James Tilley. 2021. “Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum.” *British Journal of Political Science* 51. <https://doi.org/10.1017/S0007123420000125>.
- Hopkins, Daniel J., John Sides, and Jack Citrin. 2019. “The Muted Consequences of Correct Information about Immigration.” *Journal of Politics* 81 (January): 315–20. <https://doi.org/10.1086/699914>.
- Johnson, Hollyn M., and Colleen M. Seifert. 1994. “Sources of the Continued Influence Effect: When Misinformation in Memory Affects Later Inferences.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20. <https://doi.org/10.1037/0278-7393.20.6.1420>.
- Kahan, Dan M. 2017. “Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition.” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2973067>.
- Kubin, Emily, Curtis Puryear, Chelsea Schein, and Kurt Gray. 2021. “Personal Experiences Bridge Moral and Political Divides Better Than Facts.” *Proceedings of the National Academy of Sciences of the United States of America* 118 (February). <https://doi.org/10.1073/pnas.2008389118>.
- Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F Rich. 2000. “Misinformation and the Currency of Democratic Citizenship.” *The Journal of Politics* 62: 790–816. <https://doi.org/10.1111/0022-3816.00033>.
- Kunda, Ziva. 1990. “The Case for Motivated Reasoning.” *Psychological Bulletin* 108: 480–

98. <https://doi.org/10.1037/0033-2909.108.3.480>.
- Lewandowsky, Stephan, Ullrich K. H. Ecker, Colleen M. Seifert, Norbert Schwarz, and John Cook. 2012. “Misinformation and Its Correction.” *Psychological Science in the Public Interest* 13 (September): 106–31. <https://doi.org/10.1177/1529100612451018>.
- Linden, Sander van der, Anthony Leiserowitz, Seth Rosenthal, and Edward Maibach. 2017. “Inoculating the Public Against Misinformation about Climate Change.” *Global Challenges* 1: 1600008. <https://doi.org/10.1002/gch2.201600008>.
- Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. 1st ed. Cambridge University Press.
- Lutz, Philipp, and Marco Bitschnau. 2022. “Misperceptions about Immigration: Reviewing Their Nature, Motivations and Determinants.” *British Journal of Political Science*, May, 1–16. <https://doi.org/10.1017/S0007123422000084>.
- McCright, Aaron M., Meghan Charters, Katherine Dentzman, and Thomas Dietz. 2016. “Examining the Effectiveness of Climate Change Frames in the Face of a Climate Change Denial Counter-Frame.” *Topics in Cognitive Science* 8: 76–97. <https://doi.org/10.1111/tops.12171>.
- McIntyre, Lee. 2018. *Post-Truth*. Vol. 1. The MIT Press.
- Mercier, Hugo, and Dan Sperber. 2011. “Why Do Humans Reason? Arguments for an Argumentative Theory.” *Behavioral and Brain Sciences* 34. <https://doi.org/10.1017/S0140525X10000968>.
- Mitchell, Amy, Jeffrey Gottfried, Michael Barthel, and Nami Sumida. 2018. “Distinguishing Between Factual and Opinion Statements in the News.” <https://www.pewresearch.org/journalism/2018/06/18/distinguishing-between-factual-and-opinion-statements-in-the-news/>.
- Nickell, Stephen, and Jumana Saleheen. 2017. “The Impact of EU and Non-EU Immigration on British Wages.” *IZA Journal of Development and Migration* 7: 1–28. <https://doi.org/10.1186/s40176-017-0116-0>.

- Nieminen, Sakari, and Lauri Rapeli. 2019. "Fighting Misperceptions and Doubting Journalists' Objectivity: A Review of Fact-Checking Literature." *Political Studies Review* 17: 296–309. <https://doi.org/10.1177/1478929918786852>.
- Nyhan, Brendan. 2020. "Facts and Myths about Misperceptions." *Journal of Economic Perspectives* 34 (June): 220–36. <https://doi.org/10.1257/jep.34.3.220>.
- Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32: 303–30. <https://doi.org/10.1007/s11109-010-9112-2>.
- . 2019. "The Roles of Information Deficits and Identity Threat in the Prevalence of Misperceptions." *Journal of Elections, Public Opinion and Parties* 29: 222–44. <https://doi.org/10.1080/17457289.2018.1465061>.
- Ottaviano, Gianmarco I. P., and Giovanni Peri. 2012. "Rethinking the Effect of Immigration on Wages." *Journal of the European Economic Association* 10. <https://doi.org/10.1111/j.1542-4774.2011.01052.x>.
- Parekh, Ravi, Martin Ruhs, and Vargas Silva. 2017. "How Immigrants Affect Jobs and Wages." Full Fact. <https://fullfact.org/immigration/immigration-and-jobs-labour-market-effects-immigration/>.
- Porter, Ethan, and Thomas J. Wood. 2019. *False Alarm. The Truth about Political Mistruths in the Trump Era*. Edited by Francis Lee. Cambridge University Press. 1st ed. <https://doi.org/10.1017/CBO9781107415324.004>.
- Redlawsk, David P., Andrew J. W. Civettini, and Karen M. Emmerson. 2010. "The Affective Tipping Point: Do Motivated Reasoners "Ever Get It"?" *Political Psychology* 31: 563–93. <https://doi.org/10.1111/j.1467-9221.2010.00772.x>.
- Schwartz, Cassilde, Miranda Simon, David Hudson, and Jennifer Van-Heerde-Hudson. 2021. "A Populist Paradox? How Brexit Softened Anti-Immigrant Attitudes." *British Journal of Political Science* 51. <https://doi.org/10.1017/S0007123419000656>.
- Shelby, Ashley, and Karen Ernst. 2013. "Story and Science: How Providers and Parents

- Can Utilize Storytelling to Combat Anti-Vaccine Misinformation.” *Human Vaccines and Immunotherapeutics* 9: 1795–1801. <https://doi.org/10.4161/hv.24828>.
- Shin, Jieun, and Kjerstin Thorson. 2017. “Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media.” *Journal of Communication* 67: 233–55. <https://doi.org/10.1111/jcom.12284>.
- SkyNews. 2016. “Michael Gove - ‘EU: In or Out?’ 3 June 2016.” *Sky News*. <https://www.youtube.com/watch?v=t8D8AoC-5i8>.
- Sobolewska, Maria, and Robert Ford. 2020. *Brexitland. Brexitland*. Cambridge University Press. <https://doi.org/10.1017/9781108562485>.
- Swire, Briony, Adam J. Berinsky, Stephan Lewandowsky, and Ullrich K. H. Ecker. 2017. “Processing Political Misinformation: Comprehending the Trump Phenomenon.” *Royal Society Open Science* 4: 160802. <https://doi.org/10.1098/rsos.160802>.
- Swire, Briony, and Ullrich Ecker. 2018. “Misinformation and Its Correction: Cognitive Mechanisms and Recommendations for Mass Communication.” Edited by Brian G. Southwell, Emily A. Thorson, and Laura Sheble. *Misinformation and Mass Audiences*, 195–211. <https://doi.org/10.7560/314555-013>.
- Tesich, Steve. 1992. “A Government of Lies.” *The Nation*.
- Thorson, Emily. 2016. “Belief Echoes: The Persistent Effects of Corrected Misinformation.” *Political Communication* 33: 460–80. <https://doi.org/10.1080/10584609.2015.1102187>.
- Todd, Charles. 2017. “Conway: Press Secretary Gave ‘Alternative Facts’” *NBC News*. <https://www.nbcnews.com/meet-the-press/video/conway-press-secretary-gave-alternative-facts-860142147643>.
- Vraga, Emily K., and Leticia Bode. 2017. “Using Expert Sources to Correct Health Misinformation in Social Media.” *Science Communication* 39. <https://doi.org/10.1177/1075547017731776>.
- Walter, Nathan, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag. 2020. “Fact-Checking: A Meta-Analysis of What Works and for Whom.” *Political Communication*

37. <https://doi.org/10.1080/10584609.2019.1668894>.
- Walter, Nathan, and Sheila T. Murphy. 2018. "How to Unring the Bell: A Meta-Analytic Approach to Correction of Misinformation." *Communication Monographs* 85 (July): 423–441. <https://doi.org/10.1080/03637751.2018.1467564>.
- Weeks, Brian E. 2015. "Emotions, Partisanship, and Misperceptions: How Anger and Anxiety Moderate the Effect of Partisan Bias on Susceptibility to Political Misinformation." *Journal of Communication* 65 (August): 699–719. <https://doi.org/10.1111/jcom.12164>.
- Wittenberg, Chloe, and Adam J Berinsky. 2020. "Misinformation and Its Correction." Edited by Nathaniel Persily and Joshua A. Tucker. *Social Media and Democracy The State of the Field, Prospects for Reform*. Cambridge University Press.
- Wyer, Robert S., Gerald L. Clore, and Linda M. Isbell. 1987. "Person Memory and Judgments: The Impact of Information That One Is Told to Disregard." *Advances in Experimental Social Psychology* 53: 14–29. [https://doi.org/10.1016/S0065-2601\(08\)60271-3](https://doi.org/10.1016/S0065-2601(08)60271-3).
- Ylä-Anttila, Tuukka. 2018. "Populist Knowledge: 'Post-Truth' Repertoires of Contesting Epistemic Authorities." *European Journal of Cultural and Political Sociology* 5. <https://doi.org/10.1080/23254823.2017.1414620>.
- Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge University Press. <http://www.cambridge.org/us/academic/subjects/politics-international-relations/comparative-politics/nature-and-origins-mass-opinion?format=PB&isbn=9780521407861>.

Supporting Information

Contents

Appendix A - Summary statistics	1
Appendix B - Full Results	3
OLS Models for the coefficient plots in Figure 4 and 8	6
Appendix C - Subgroup Analyses	9
Effect of the fact-check and the undermining message among respondents who were motivated to reject the statistics	9
Appendix C - Pilot Study	14
Appendix D - Questionnaire	17

Appendix A - Summary statistics

Table 1: Demographics

	No of asylum seekers (N=1176)	Cost of immigration (N=702)	Crime in London (N=480)	Effect on low wages (N=578)
Gender				
N-Miss	2	0	6	3
Male	526 (45%)	298 (42%)	239 (50%)	271 (47%)
Female	648 (55%)	404 (58%)	235 (50%)	304 (53%)
Age				
Mean	47	46	39	42
Range	18 - 88	18 - 99	18 - 81	18 - 99
Region				
South	397 (34%)	229 (33%)	145 (30%)	182 (31%)
North	292 (25%)	163 (23%)	123 (26%)	150 (26%)
London	148 (13%)	86 (12%)	78 (16%)	61 (11%)
Midlands	175 (15%)	126 (18%)	78 (16%)	103 (18%)
Scotland	108 (9%)	51 (7%)	42 (9%)	50 (9%)
Wales	56 (5%)	47 (7%)	14 (3%)	32 (6%)
Left/Right				
Ideology				
N-Miss	246	180	79	93
left-wing	243 (26%)	117 (22%)	196 (49%)	214 (44%)
centre	398 (43%)	275 (53%)	141 (35%)	194 (40%)
right-wing	289 (31%)	130 (25%)	64 (16%)	77 (16%)

	No of asylum seekers (N=1176)	Cost of immigration (N=702)	Crime in London (N=480)	Effect on low wages (N=578)
Immigration				
Preferences				
fewer	709 (60%)	447 (64%)	103 (21%)	129 (22%)
no change	306 (26%)	159 (23%)	211 (44%)	263 (46%)
more	161 (14%)	96 (14%)	166 (35%)	186 (32%)

Table 2: Treatment Groups

	No of asylum seekers (N=1176)	Cost of immigration (N=702)	Crime in London (N=480)	Effect on low wages (N=578)
Treatment Group				
No undermining message	285 (24%)	181 (26%)	137 (29%)	134 (23%)
Professor: Expert bias	155 (13%)	80 (11%)	58 (12%)	64 (11%)
Professor: Personal experience	135 (11%)	100 (14%)	55 (11%)	71 (12%)
Professor: Truthiness	152 (13%)	97 (14%)	54 (11%)	85 (15%)
Blogger: Expert bias	138 (12%)	95 (14%)	61 (13%)	67 (12%)
Blogger: Personal experience	164 (14%)	82 (12%)	55 (11%)	74 (13%)
Blogger: Truthiness	147 (12%)	67 (10%)	60 (12%)	83 (14%)

Appendix B - Full Results

Figure 9 shows how the four fact-checks affected belief each of the four false claims. To simplify comparisons we use the length of a line to illustrate how far respondents moved in their accuracy perceptions after they saw a fact-check. The grey dots show how true or false they rated the respective false claim when they first saw it at the beginning of the survey. The orange dots show how true or false they rated the respective false claim at the end of the survey, after they had seen a fact-check (and, if treated, an undermining message). Accuracy ratings are shown on a seven-point scale from 0 (Definitely true) to 6 (Definitely false).³ The grey lines connecting the grey and orange dots show how much closer they moved to the ‘false’ end of the scale. The longer the line the more respondents changed their mind after seeing the statistics.

The pattern is similar: For each line, the top line (showing those who saw the additional undermining message) is shorter than the bottom line (showing those who saw the fact-check only). In other words, the undermining message always cancelled out some of the effect of the fact-check. *How much* it cancelled out was different for each comment:

The first panel shows respondents who rated the false statement that “There has been a sharp rise in the number of people applying for asylum in the UK in the past ten years” closest to true and, therefore, were directed to a fact-check showing statistics on the number of asylum seekers coming into the UK. This fact-check shifted average accuracy ratings 1.98 points closer to the ‘false’ end of the scale (from 1.1 to 3.08) or, if accompanied by an undermining message, 1.38 points closer to ‘false’ (from 0.97 to 2.35). Overall, then, the undermining message cancelled out 30% ($1-1.38/1.98$) of the effect of the fact-check about the number of asylum seekers.

The second panel shows respondents who rated the false statement that “Immigrants receive

³Note that the original scale (as shown in the questionnaire) ranged from 0 (Definitely false) to 6 (Definitely true). For ease of interpretation we re-coded it as 0 (Definitely true) to 6 (Definitely false).)

more in benefits and services than they pay in taxes” closest to true and, therefore, were directed to a fact-check showing statistics on the net fiscal contributions of European immigrants and native British citizens. This fact-check shifted average accuracy ratings 1.61 points closer to the ‘false’ end of the scale (from 1.32 to 2.93) or, if accompanied by an undermining message, 1.2 points closer to ‘false’ (from 1.46 to 2.67). Overall, then, the undermining message cancelled out 25% ($1-1.2/1.6$) of the effect of the fact-check about the number of asylum seekers.

The third panel shows respondents who rated the false statement that “The majority of crimes in London are committed by white people, not ethnic minorities” closest to true and, therefore, were directed to a fact-check showing statistics on minority populations and minority proportions for offenders and victims of various types of reported crime in London. This fact-check shifted average accuracy ratings 1.61 points closer to the ‘false’ end of the scale (from 1.32 to 2.93) or, if accompanied by an undermining message, 1.20 points closer to ‘false’ (from 1.46 to 2.67). Overall, then, the undermining message cancelled out 25% ($1-1.2/1.61$) of the effect of the fact-check about crime in London.

The fourth panel shows respondents who rated the false statement that “Immigration to the UK does not affect the wages of the low-paid” closest to true and, therefore, were directed to a fact-check showing statistics on the effect of immigration on hourly wages for workers employed in the semi- and unskilled services in the UK. This fact-check shifted average accuracy ratings 1.4 points closer to the ‘false’ end of the scale (from 1.19 to 2.59) or, if accompanied by an undermining message, 1.24 points closer to ‘false’ (from 1.21 to 2.45). Overall, then, the undermining message cancelled out 11% ($1-1.24/1.4$) of the effect of the fact-check about the number of asylum seekers.

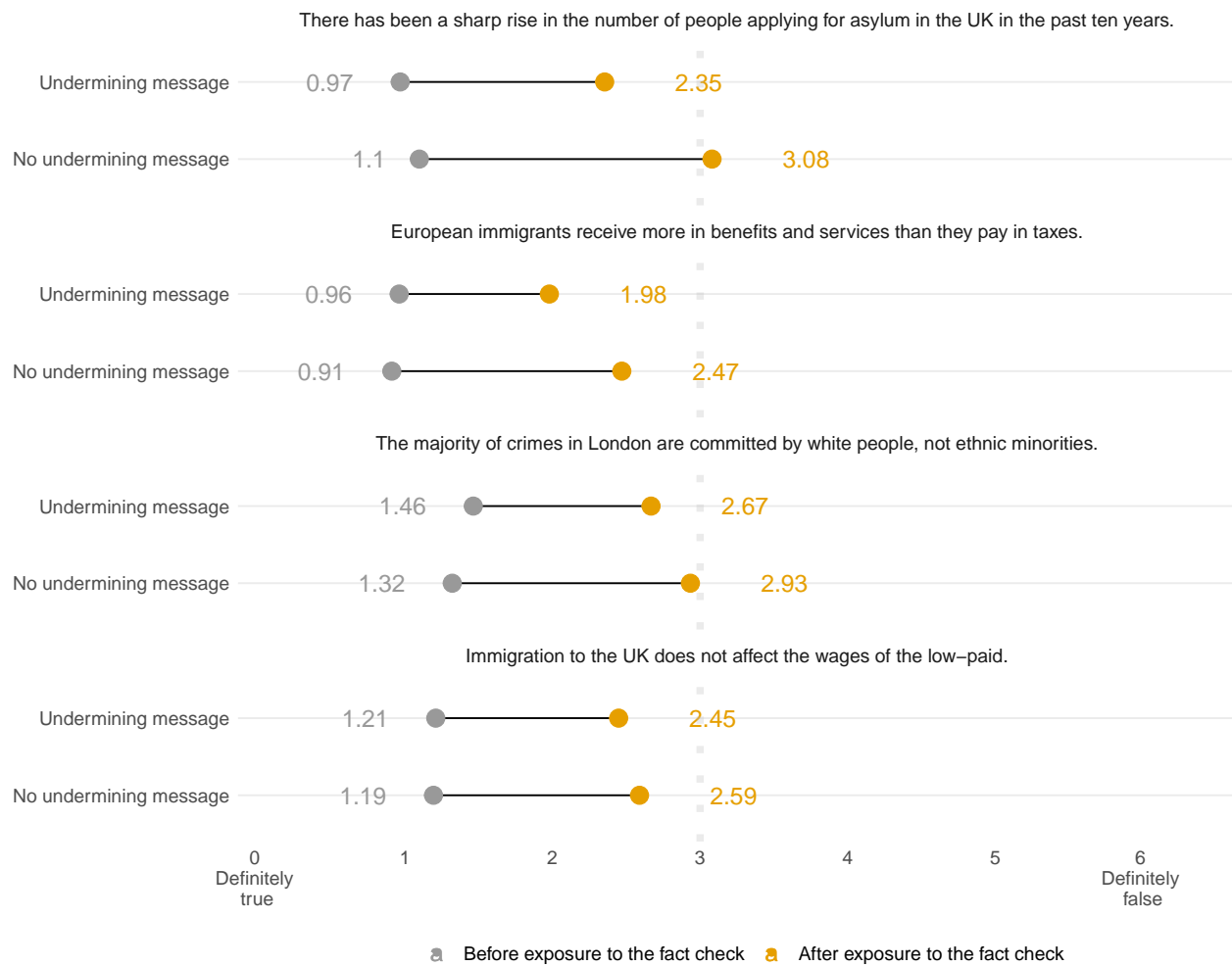


Figure 9: Average belief in false facts before and after exposure to a fact-check.

OLS Models for the coefficient plots in Figure 4 and 8

Table 3 shows the three OLS regression models underlying the coefficient plot in Figure 4 of the main body of the text. The models estimate the effect of exposure to a fact-check on three indicators of a successful fact-check: the difference in the seven-point ‘true’ to ‘false’ ratings before and after exposure to the respective fact check (T2-T1), the perceived accuracy of the expert’s information (1 not at all accurate to 4 very accurate), and trust in what the expert says on immigration (1 not at all accurate to 4 very accurate).

Table 4 shows the three OLS regression models underlying the coefficient plot in Figure 8 of the main body of the text. The models estimate the effect of exposure to a fact-check on three indicators of post-truth reasoning: fact-opinion ratings, agreement that ‘It’s OK to disagree with the facts if that’s what you believe’, and the probability of choosing the statement that “The statistics are probably right, but I believe something different” to explain where respondents stood. Fact-opinion ratings are measured on a scale from 0 (Purely a matter of fact) to 6 (Purely a matter of opinion). The willingness to disagree with the facts is measured on a scale from 1 (Strongly disagree) to 4 (Strongly agree).

Table 3: Effect of the six post-truth comments on belief in the false claims, and trust in the expert, and their information

	<i>Dependent variable:</i>		
	difference in belief	perceived accuracy	trust in expert
	OLS (1)	OLS (2)	OLS (3)
Prof: Expert bias	−0.52*** (0.12)	−0.08* (0.04)	−0.21** (0.09)
Prof: Personal Experience	−0.30*** (0.11)	−0.12*** (0.04)	−0.18** (0.09)
Prof: Truthiness	−0.43*** (0.11)	−0.07 (0.04)	−0.05 (0.09)
Blogger: Expert bias	−0.54*** (0.11)	−0.09** (0.04)	−0.29*** (0.09)
Blogger: Personal Experience	−0.52*** (0.11)	−0.10** (0.04)	−0.20** (0.09)
Blogger: Truthiness	−0.47*** (0.12)	−0.08* (0.04)	−0.14 (0.09)
Observations	2,936	2,936	2,936
R ²	0.01	0.004	0.005
Adjusted R ²	0.01	0.002	0.003
Residual Std. Error (df = 2929)	1.78	0.69	1.45

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4: Effect of the six post-truth comments on three indicators of post-truth reasoning

	<i>Dependent variable:</i>		
	Matter of opinion	Okay to disagree	Believe something different
	OLS	OLS	Logit
	(1)	(2)	(3)
Prof: Expert bias	0.06 (0.11)	0.03 (0.05)	0.02 (0.04)
Prof: Personal Experience	0.17 (0.11)	0.08 (0.05)	-0.02 (0.04)
Prof: Truthiness	0.05 (0.11)	0.04 (0.05)	-0.13*** (0.04)
Blogger: Expert bias	0.02 (0.11)	0.04 (0.05)	-0.03 (0.04)
Blogger: Personal Experience	0.12 (0.11)	0.01 (0.05)	-0.04 (0.04)
Blogger: Truthiness	-0.02 (0.11)	0.05 (0.05)	-0.06 (0.04)
Observations	2,936	2,936	1,337
R ²	0.001	0.001	0.01
Adjusted R ²	-0.001	-0.001	0.01
Residual Std. Error	1.78 (df = 2929)	0.82 (df = 2929)	0.45 (df = 1330)

Note:

*p<0.1; **p<0.05; ***p<0.01

Appendix C - Subgroup Analyses

Effect of the fact-check and the undermining message among respondents who were motivated to reject the statistics

Up to here, our analyses are based on the entire sample, including some who did not see immigration as an important issue and some whose most confident false belief ran against the grain of their opinions on immigration. For instance, some of those who said that Britain should take in *more* immigrants (so, pro-immigration respondents) thought, incorrectly, that the number of asylum seekers in the UK were rising, and so were directed to statistics proving them wrong. These respondents should be happy to be proven wrong. On the hand, some of those who said that Britain should take in *fewer* immigrants (so, anti-immigration respondents) thought, incorrectly, that immigration to the UK had no effect on low-paid wages. These respondents will have been similarly happy to read the fact-check.

To examine the effect of the fact-check and the undermining message on those who had a motivation to reject the statistics we created a subsample removing those these respondents (n=882, or 30% of the sample). The following analyses are based on a subset of respondents who said that immigration was at least ‘somewhat important’ and who were directed to a fact-check that challenged not only their factual beliefs but also their immigration preferences.

The bar plot in Figure 10 (analogous to Figure 2 in the main text) shows the effect of the fact-check and the undermining message on belief in false claims among this subsample of respondents who were motivated to reject the statistics (n=2054). The fact-check alone shifted perceptions 1.67 points along the scale (from 0.91 to 2.58). The fact-check followed by an undermining message shifted perceptions a mere 1.19 points along the scale (from 0.92 to 2.11). That means that the offsetting effect of the undermining message rises to 29% ($1-1.19/1.67$) among those who had a motivation to reject the fact-check. Figure 11

(analogous to Figure 9 for the whole sample) illustrates the effect of the fact-check and the undermining message for each fact-check group.

The OLS and logit models in tables 5 and 6 show the effect of the treatments on the additional outcome variables among this subsample. Here, too the effects are very similar to the effects in the whole sample (tables 3 and 4). The finding that reactions to the undermining message did not really depend on respondent's incentives to reject the statistics and be nudged back to their prior beliefs is noteworthy: It means that factual beliefs drive responses to fact-checks much more than partisan motivations. How convinced people are the a statement is true is a more powerful predictor than how much they want it be true.

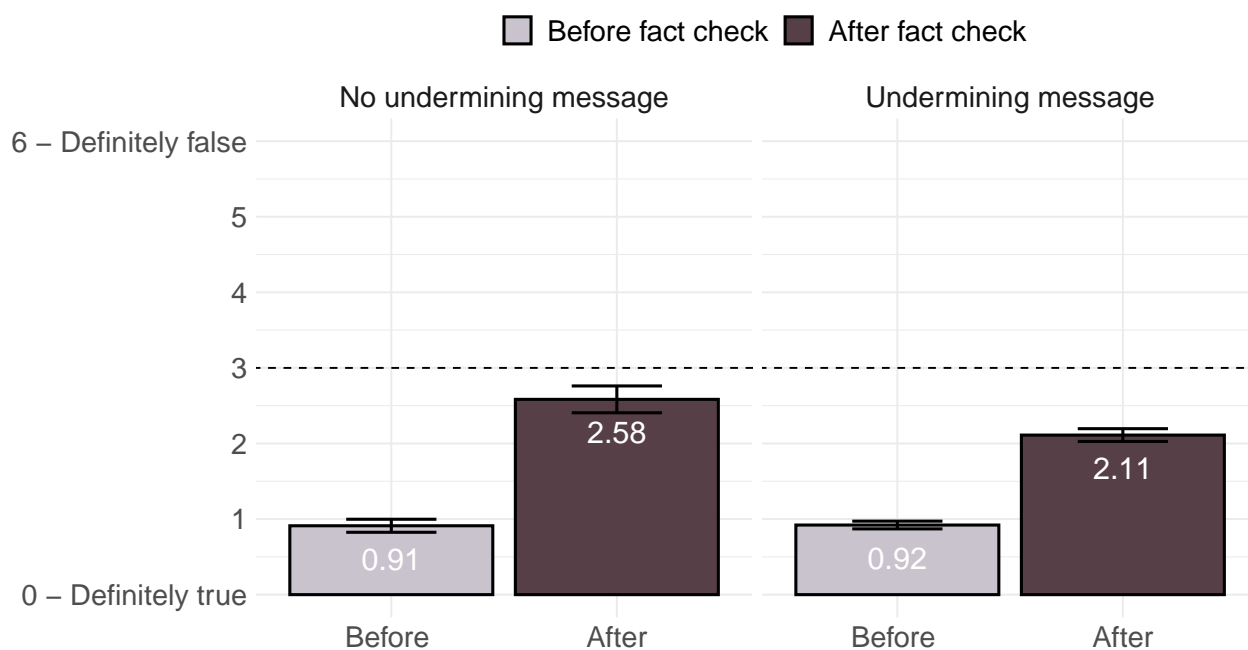


Figure 10: Average assessment of the false claims by exposure to an undermining message. Subset of respondents who had an incentive to disregard the fact-check they saw.

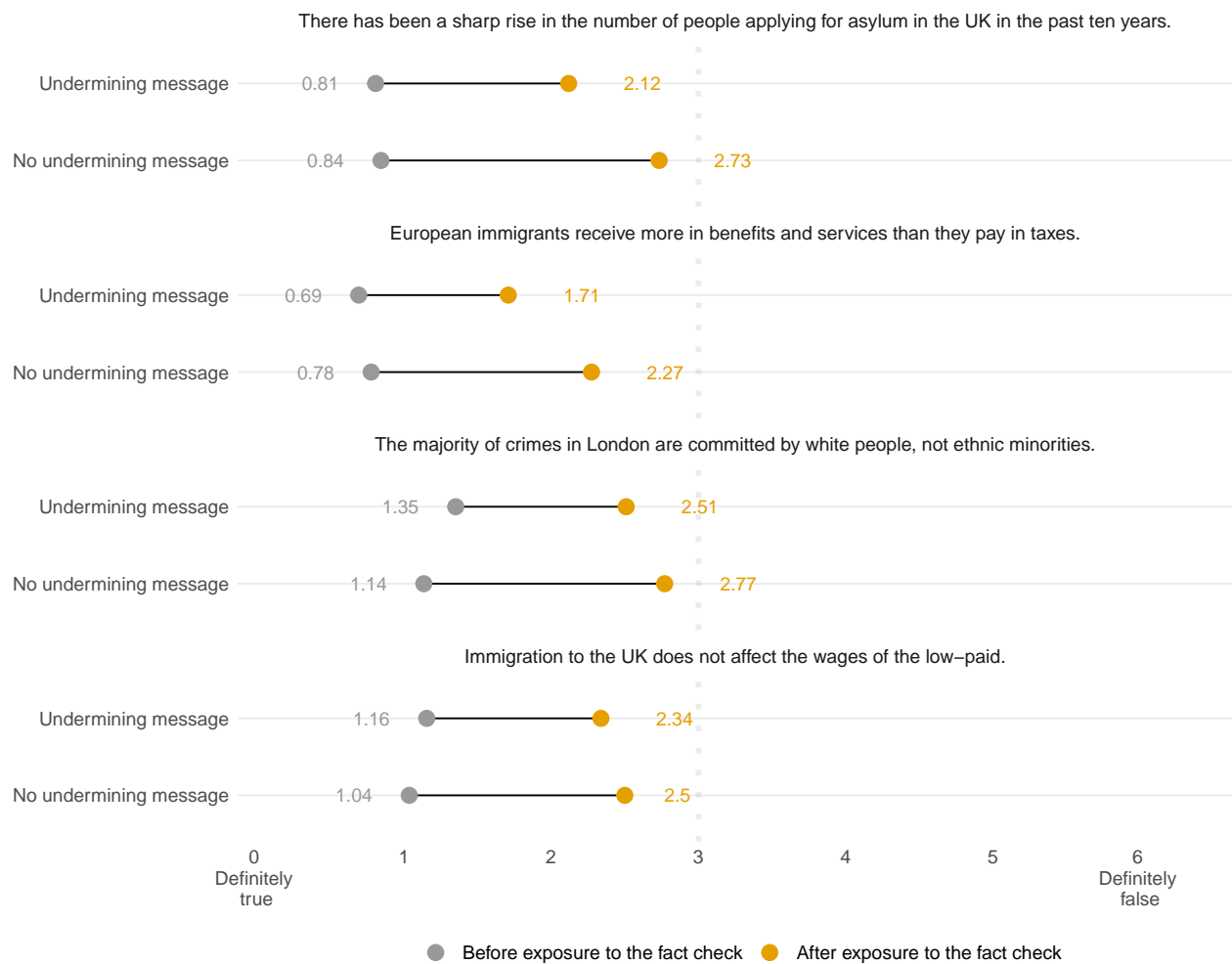


Figure 11: Respondents who were motivated to reject the statistics: Average belief in false facts before and after exposure to a fact-check.

Table 5: Respondents who were motivated to reject the statistics: Effect of the six post-truth comments on belief in the false claims, and trust in the expert, and their information

	<i>Dependent variable:</i>		
	difference in belief	perceived accuracy	trust in expert
	OLS	OLS	OLS
	(1)	(2)	(3)
Prof: Expert bias	−0.45*** (0.14)	−0.08 (0.05)	−0.25** (0.12)
Prof: Personal Experience	−0.32** (0.14)	−0.16*** (0.05)	−0.30*** (0.12)
Prof: Truthiness	−0.48*** (0.13)	−0.09* (0.05)	−0.14 (0.11)
Blogger: Expert bias	−0.57*** (0.14)	−0.09 (0.05)	−0.33*** (0.12)
Blogger: Personal Experience	−0.57*** (0.14)	−0.08 (0.05)	−0.19 (0.12)
Blogger: Truthiness	−0.50*** (0.14)	−0.12** (0.05)	−0.22* (0.12)
Observations	2,054	2,054	2,054
R ²	0.02	0.01	0.01
Adjusted R ²	0.01	0.003	0.003
Residual Std. Error (df = 2047)	1.78	0.70	1.52

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 6: Respondents who were motivated to reject the statistics: Effect of the six post-truth comments on three indicators of post-truth reasoning

	<i>Dependent variable:</i>		
	Matter of opinion	Okay to disagree	Believe something different
	OLS	OLS	Logit
	(1)	(2)	(3)
Prof: Expert bias	0.12 (0.14)	0.03 (0.06)	-0.01 (0.05)
Prof: Personal Experience	0.16 (0.14)	0.07 (0.06)	-0.06 (0.05)
Prof: Truthiness	0.05 (0.14)	0.04 (0.06)	-0.17*** (0.05)
Blogger: Expert bias	0.02 (0.14)	0.06 (0.06)	-0.08 (0.05)
Blogger: Personal Experience	0.07 (0.14)	0.05 (0.06)	-0.07 (0.05)
Blogger: Truthiness	-0.06 (0.14)	0.04 (0.06)	-0.08 (0.05)
Observations	2,054	2,054	926
R ²	0.001	0.001	0.01
Adjusted R ²	-0.002	-0.002	0.01
Residual Std. Error	1.81 (df = 2047)	0.82 (df = 2047)	0.46 (df = 919)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 7: Trust in information from various sources (Pilot Study 1)

Immigration preferences Variable	anti-immig			pro-immig		
	N	Mean	SD	N	Mean	SD
Journalist from the BBC News Reality Check Team	54	2.7	0.88	69	2.9	1.1
Researcher at the at the House of Commons Library	54	3	1	69	3	0.99
Research fellow at the Social Market Foundation	55	2.7	0.91	69	3	0.98
UK Border Agency official	55	3.6	1	69	3.2	1
Spokesperson at the Office for National Statistics	55	3.5	0.94	69	3.7	0.92
Researcher at Migration Watch UK	55	3.3	0.91	69	3.1	0.96
Professor of Economics at the University of Oxford	54	3.1	0.9	69	3.3	0.94
Fellow at Chatham House	55	3.1	0.88	68	3.3	0.84
Researcher at Oxford University’s Migration Observatory	55	3.3	0.86	69	3.6	0.86

Note: Pilot 1 (24 Nov 2018), prolific academic, n=124 UK respondents

Appendix C - Pilot Study

We conducted two small pilot studies to identify prominent false beliefs about immigration, and a trustworthy source for our fact checks. In the first pilot fielded on 24 Nov 2018, we tried to identify a source for the fact-checks that would be equally trusted by respondents with different views on immigration. Table 7 shows average levels of trust in information about immigration from various sources on a scale from 1 (not at all) to 5 (completely). The most trusted source among respondents on both sides of the immigration divide was the ‘spokesperson at the Office for National Statistics’; followed by the ‘Professor of Economics at the University of Oxford’. Hence, we made our fact-checker a Professor of Economics at the University of Oxford and gave him work experience at the Office for National Statistics. To keep the authoritative version of the undermining message at about the same level of authoritativeness as the fact-check we also attributed it to a ‘Professor of Economics’, switching the institution to the London School of Economics.

To identify common misperceptions we also piloted a number of false facts from from UK fact checking websites. In our first pilot study, we asked respondents to rate the following statements as ‘true’ or ‘false’: (1) ‘There has seen a sharp rise in the number of people

Table 8: Belief in false statements (Pilot Study 1)

Immigration preferences Variable	anti-immig			pro-immig		
	N	Mean	SD	N	Mean	SD
(1) sharp rise in number of asylum seekers	54	0.93	0.26	68	0.74	0.44
(2) UK takes in its fair share	55	0.65	0.48	69	0.32	0.47
(3) 10 per cent Muslim	55	0.71	0.46	69	0.43	0.5
(4) immigrants receive more than they pay	55	0.55	0.5	69	0.23	0.43
(5) no effect on low wages	55	0.38	0.49	69	0.7	0.46
(6) courts more likely to convict black defendants	55	0.4	0.49	69	0.64	0.48
(7) UK includes students in migration statistics	55	0.38	0.49	69	0.39	0.49
(8) net migration has not increased	55	0.33	0.47	69	0.45	0.5

Note: Pilot 1 (24 Nov 2018), prolific academic, n=124 UK respondents.

applying for asylum in the UK in the past 10 years.’, (2) ‘Compared to countries like France, Spain and the Netherlands, the UK takes more than its fair share of refugees.’, (3) ‘More than 10 per cent of the UK population is Muslim.’, (4) ‘Immigrants receive more in benefits and services than they pay in taxes.’, (5) ‘Immigration to the UK does not affect the wages of the low paid.’, (6) ‘UK courts are more likely to convict black defendants than white defendants.’, (7) ‘The UK is the only major Western democracy to include international students in its net migration statistics.’, (8) ‘The rate of net migration to the UK has not increased much in recent decades.’ We found high levels of belief in three of the anti-immigration statements but relatively low levels of belief in the pro-immigration statements.

In our second pilot study, we tested belief in a few more false statements that we thought might be popular among pro-immigration respondents. We replaced the binary true/false scale with a four-point scale, asking respondents to rate each false statement as ‘definitely true’ (4), ‘probably true’ (3), ‘probably false’ (2), or ‘definitely false’ (1). Respondents assessed the following statements: (1) ‘There has been a sharp rise in the number of people applying for asylum in the UK in the past 10 years.’; (2) ‘Immigrants in the UK receive more in benefits and services than they pay in taxes.’; (3) ‘Immigration from Europe has increased crime.’; (4) ‘In proportion to its population the UK takes more refugees than

Table 9: Belief in false facts (Pilot Study 2)

Variable	Immigration preferences			anti-immig			pro-immig		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
(1) sharp rise in the number of asylum seekers	62	3.1	0.7	90	2.9	0.66			
(2) immigrants receive more than they pay	41	3.1	0.83	52	1.9	0.94			
(3) European immigration has increased crime	55	3.1	0.76	75	2.2	0.85			
(4) UK takes in its fair share	30	2.8	0.85	59	2.3	0.81			
(5) most London crimes committed by white people	47	2.5	0.78	52	2.7	0.66			
(6) no effect on low wages	40	1.8	0.81	47	2.7	0.68			
(7) courts more likely to convict black defendants	63	2.2	0.95	93	2.7	0.87			
(8) proportion of immigrants in prison	64	2.3	0.7	91	2.2	0.76			
(9) no increase in net migration	63	2	0.99	86	2.3	0.76			

Note: 29 Nov 2018, prolific academic, n=199 UK respondents.

nearby countries like France, Denmark and the Netherlands.’; (5) ‘The majority of crimes in London are committed by white people.’; (6) ‘Immigration to the UK does not affect the wages of the low-paid.’; ⁴ (7) ‘UK courts are more likely to convict black defendants than white defendants.’; (8) ‘The proportion of immigrants in prison is approximately the same as the proportion of immigrants in the UK.’; (9) ‘The rate of net migration to the UK has not increased much in recent decades.’ For the main study, we chose four claims that were clearly false and easily proven wrong – two that were popular among those with anti-immigration views (statements 1 and 2), and two that were popular among those with pro-immigration views (statements 5 and 6).

⁴U.S. readers may be surprised to find statement 6 (“Immigration to the UK does not affect the wages of the low-paid”) categorized as false. There is an ongoing debate about the overall effect of immigration on local wages, with some U.S. studies (most notably, Ottaviano and Peri (2012)) finding a positive effect on local wages. In the U.K., however, recent research has shown that the effect differs of immigration on wages depends on the sector: Immigration to the U.K. has had a positive, or no effect on wages for high-skilled jobs, but a negative effect on wages for low-skilled jobs (Parekh, Ruhs, and Silva 2017; Nickell and Saleheen 2017; Dustmann, Frattini, and Preston 2013).

Appendix D - Questionnaire

Stage 1

Background characteristics

Q1 [**gender**] Which of the following best describes how you think of yourself? [OPTIONS: Male; Female; In another way]

Q2 [**age**] And how old are you?

Q3 [**region**] Where do you live? [OPTIONS: Northern Ireland; Scotland; North West; North East; Yorkshire and the Humber; Wales; West Midlands; East Midlands; South West; South East; Eastern; London; Other]

Q4 [**education**] What is the highest educational level that you have achieved?

[OPTIONS: Secondary school; University degree or equivalent professional qualification, NVQ level 4 etc; Higher university degree, doctorate, MBA, NVQ level 5 etc; Still in full time education; No formal education; Don't know/prefer not to answer]

Q5 [**profession**] Please indicate which one of the following best describes the profession of the chief income earner in your household.

[OPTIONS: High managerial, administrative or professional e.g. doctor, lawyer, medium / large company director (50+ people); Intermediate managerial, administrative or professional e.g. teacher, manager, accountant; Supervisor, administrative or professional e.g. policeman, nurse, secretary, self-employed; Skilled manual worker e.g. mechanic, plumber, electrician, lorry driver, train driver; Semi-skilled or unskilled manual worker e.g. waiter, factory worker, receptionist, labourer; House-wife / house-husband; Unemployed; Student; Retired]

Q6 [**postcode**] This survey needs you to enter your full postcode before proceeding. The postcode is used purely to collect information about the area you live in to be analysed in relation to the other data we collect. It will not be used to identify you individually, or be

used for any other purpose other than for academic research. We would like to reassure you that your answers are anonymous and will be completely confidential. PLEASE WRITE IN YOUR FULL POSTCODE:

Q7 [**voteTom**] On a scale of 0 (certain NOT to vote) to 10 (absolutely certain to vote), how likely would you be to vote in a general election tomorrow?

Q8 [**voteTomParty**] If there were a general election held tomorrow, which party would you vote for?

[OPTIONS: Conservative; Labour; Liberal Democrat; UK Independence Party (UKIP); Scottish National Party (SNP); Plaid Cymru (PC); Green; Some other party; Would not vote; Don't know; Refuse]

Q9 [**voteTomParty2**] And now imagine The Independent Group (TIG) and The Brexit Party put up candidates at the next general election. The Conservative party, Labour, Liberal Democrats, UKIP and other parties also stand. How would you then vote? ...

Q10 [**vote2017**] Talking to people about the General Election on June 8th 2017, we have found that a lot of people didn't manage to vote. How about you - did you manage to vote in the General Election in 2017? [OPTIONS: Yes; No]

Q11 [**vote2017Party**] Which party did you vote for in the General Election on June 8th 2017? ...

Q12 [**attention**] How much attention do you generally pay to politics? Please use the following scale, where 0 means no attention at all and 10 means a great deal of attention

Q13 [**issues**] Here are some issues facing Britain today. Please choose the three issues that are most important to you and rank them from 1 (most important) to 3 (third most important). *[randomize order]*

[OPTIONS: Immigration [issues_immig]; The economy [issues_econ]; NHS/Health [issues_nhs]; Crime [issues_crime]; EU/Brexit [issues_eu]; Housing [issues_housing];

Schools/Education

Here is a short list of statements. How much you agree or disagree with each one? *[randomize order]*

Q14 [**ethno1**] Britain has a lot to learn from other countries in running its affairs.

Q15 [**pop_no_compromise**] What people call ‘compromise’ in politics is really just selling out on one’s principles.

Q16 [**ethno2**] I would rather be a citizen of Britain than of any other country in the world.

Q17 [**pop_people_make_decisions**] The people, and not politicians, should make our most important policy decisions.

Q18 [**immigFeelLikeHome**] There are so many foreigners round here that it doesn’t feel like home any more.

Q19 [**ethno6**] I am often less proud of Britain than I would like to be.

[OPTIONS: Strongly Agree, Agree, Disagree, Strongly Disagree, Don’t Know]

Q20 [**rightLeft**] Some people talk about ‘left’, ‘right’, and ‘centre’ to describe parties and politicians. Where would you place yourself on this scale?

[OPTIONS: Very left wing, fairly left wing, slightly left wing, centre, slightly right wing, fairly right wing, very right wing, don’t know]

Q21 [**immig**] Some people think that we should allow many fewer immigrants come to the UK and others think that we should allow many more immigrants. Where would you place yourself on this scale?

[OPTIONS: -3 = many fewer, 0 = no change, +3 = many more]

Q22 [**immigImportance**] How important to you is this issue?

[OPTIONS: Extremely important (4), very important (3), somewhat important (2), not at all important (1)]

Now here is another list of statements. Again, please say how much you agree or disagree.
[randomize order]

Q23 [**ethno3**] There are some things about Britain today that make me ashamed to be British.

Q24 [**ethno5**] The world would be a better place if people from other countries were more like the British.

Q25 [**pop_rep_by_citizen**] I would rather be represented by a citizen than by a specialized politician.

Q26 [**pop_pol_talk_too_much**] Elected officials talk too much and take too little action. Q27 [**immigEcon**] Immigration is good for the British economy.

Q28 [**ethno4**] People in Britain are too ready to criticise their country.

[OPTIONS: Strongly Agree; Agree; Disagree; Strongly Disagree; Don't Know]

Belief in false claims about immigration

And here is a third list of statements. This time, we'd like you to say whether, to the best of your knowledge, they are true or false? *[randomize order]*

Q29 [**noAsylum_T1**] There has been a sharp rise in the number of people applying for asylum in the UK in the past ten years.

Q30 [**costImmig_T1**] European immigrants receive more in benefits and services than they pay in taxes.

Q31 [**whiteCrime_T1**] The majority of crimes in London are committed by white people, not ethnic minorities.

Q32 [**lowPaid_T1**] Immigration to the UK does not affect the wages of the low-paid.

Q33 [**no5_T1**] Britain is the fifth largest economy in the world.

Q34 [**recession_T1**] The Leave vote in June 2016 did not result in an instant UK recession.

Q35 [**fracking_T1**] Fracking causes earthquakes.

Q36 [plasticBags__T1] England's plastic bag usage dropped 85% since 5p charge was introduced.

[OPTIONS: 0 Definitely False; 1; 2; 3; 4; 5; 6 Definitely True]

Stage 2 - Fact-Check

We have asked Richard Clarke, Professor of Economics at the University of Oxford to provide us with information about the statements you just read. On the following page, you will see some detailed information about one of these statements, that is:

[Re-print the false statement respondents rated closest to true, i.e. Q29, Q30, Q31, or Q32. If two or more were given the same highest score, randomly assign respondents to one of them. Then show the respective fact-check.]



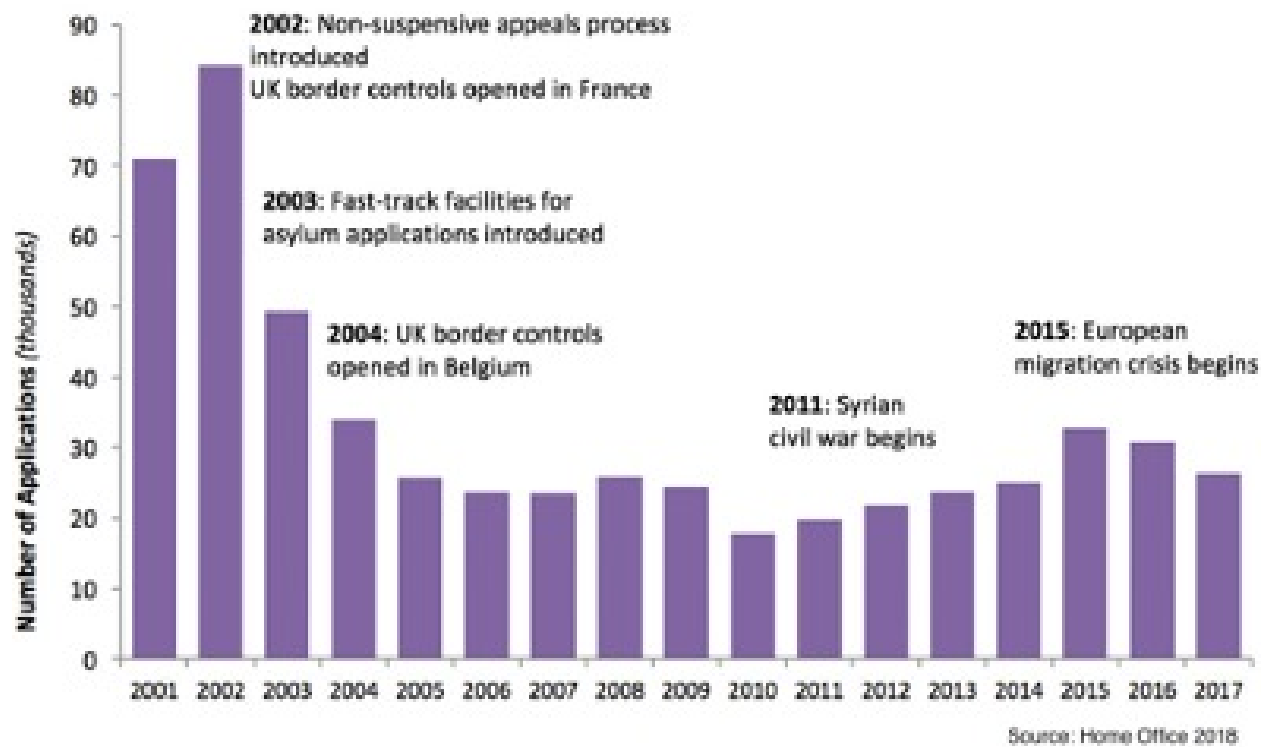
By Professor Richard Clark

University of Oxford

Professor Richard Clark is a Professor of Population Economics at the University of Oxford. From 2009 to 2015, he was a Consultant on Population and Demography with the Office for National Statistics (ONS). Professor Clarke holds a PhD from Harvard University.

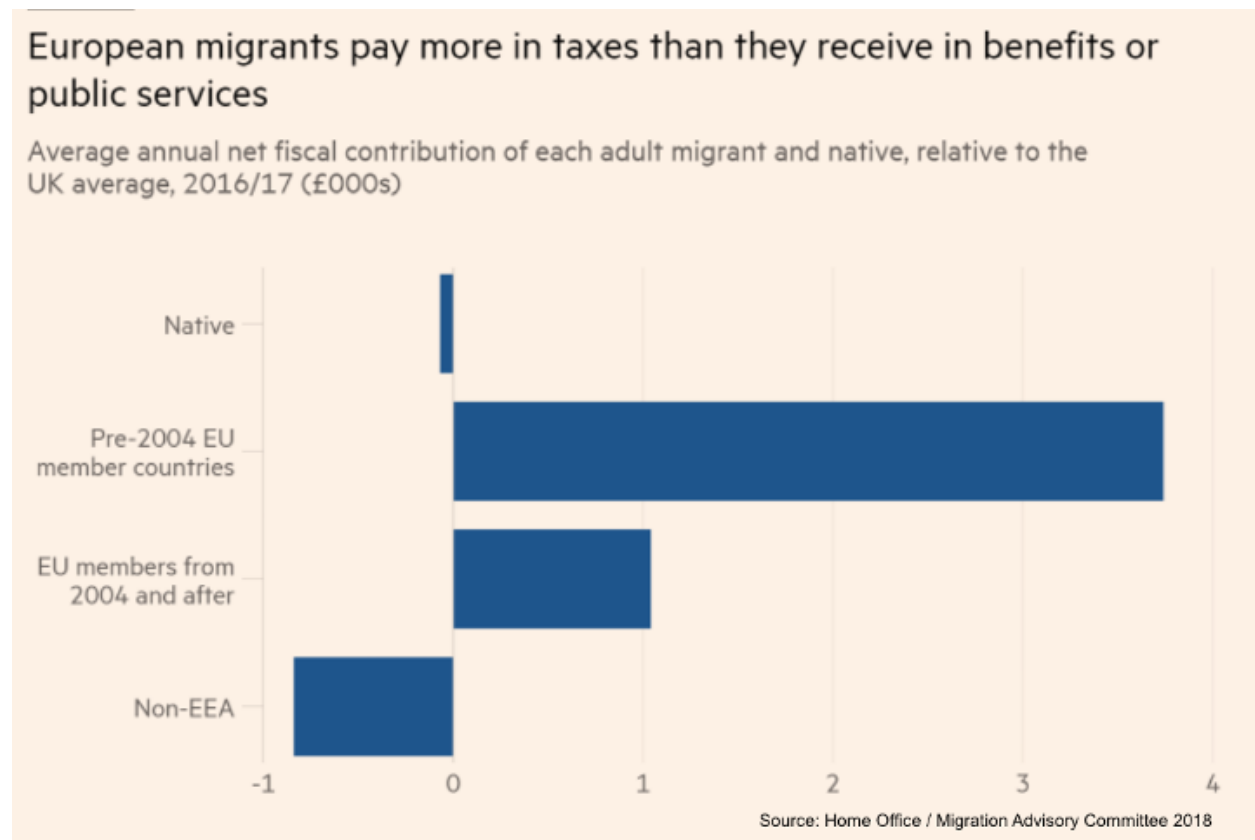
[Insert respective fact-check, as shown below]

Fact-check: “There has been a sharp rise in the number of people applying for asylum in the UK in the past 10 years.”



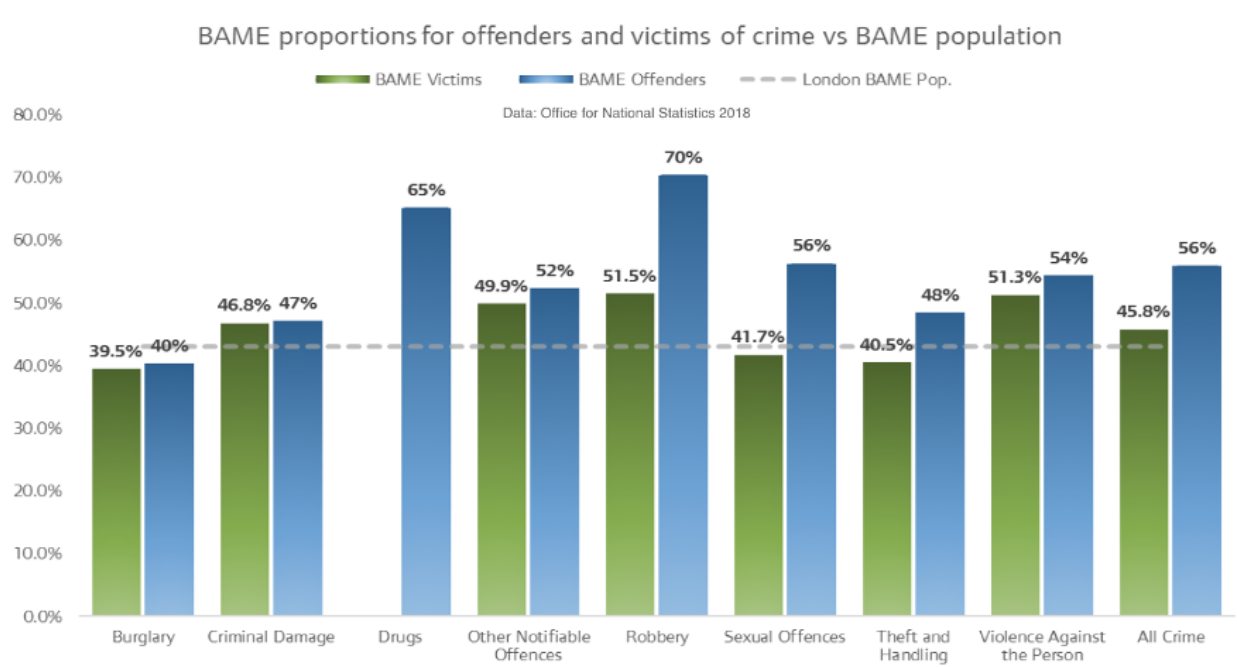
This statement is false. Asylum applications in the UK have remained stable over the past ten years. It is true that there has been a sharp rise in the number of people crossing into Europe in 2015. Few of these made it to the UK. Germany, Sweden and Hungary shouldered most of the asylum claims. The graph below is based on the official immigration statistics that the Home Office publishes every year. It shows the number of asylum applications the UK received since 2001. The UK receives about the same number of refugees today than it did ten years ago.

Fact-check: “European immigrants receive more in benefits and services than they pay in taxes.”



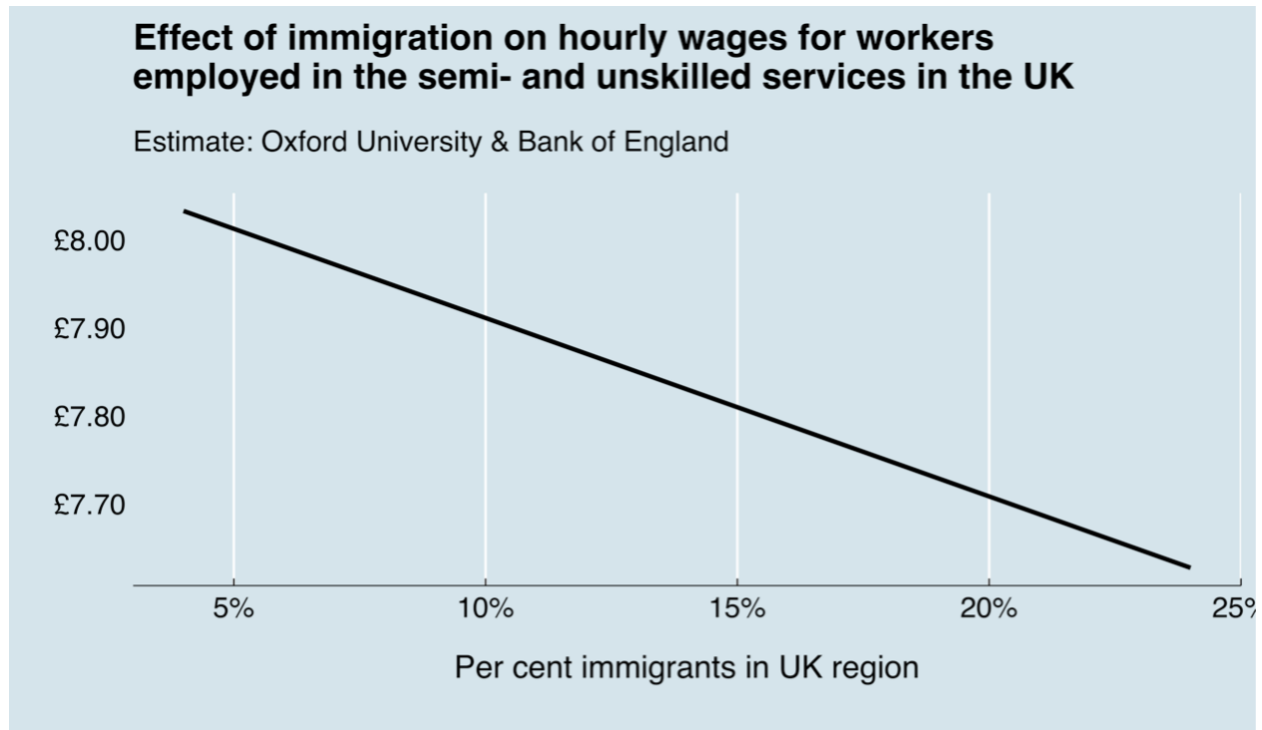
This statement is false. European immigrants pay substantially more in taxes to the government than they receive in benefits or public services. According to the most recent data European migrants made a total contribution of £4.7bn to the public finances in 2016-17. This graph shows how much migrants contribute relative to how much the average UK citizen contributes. An average adult migrant from one of the 13 countries that joined the EU before 2004 (most of Western Europe) contributed £3,740 more to Britain’s exchequer than an average UK citizen. Migrants from countries that joined after 2004 (that includes Eastern European countries) paid an average of £1,040 more in taxes than they receive.

Fact-check: “The majority of crimes in London are committed by white people.”



This statement is false. According to data from the Office for National Statistics (ONS) 43% of the population in London in 2017 identified as black, Asian and minority ethnic (BAME). But, as on the right-hand side of the graph below, 56% of all crime offenders were BAME. How overrepresented BAME offenders are depends on the type of crime. Every blue bar that exceeds the 43% dotted line means that, for this crime, BAME people are overrepresented. There are three offenses for which BAME people are particularly overrepresented: Drugs, Robbery, and Sexual Offences.

Fact-check: “Immigration to the UK does not affect the wages of the low-paid.”



This statement is false. You have to distinguish between sectors. It is true that in many sectors immigration has no effect on wages. But things are different in the less skilled sector – that is, for people working as cleaners, in care homes, in bars and restaurants, and so on. One of my Oxford colleagues and his co-author at the Bank of England looked at the effect of immigration on wages in different sectors from 1992 to 2016. Their results are clear: in the semi- and unskilled sector, 10 per cent more immigrants in a region led to a 2.6 per cent reduction in hourly pay.

[Buffer questions]

Q37 [**marital**] Which of the following best describes your marital status? [OPTIONS: Single; Married; Civil partnership; Co-habiting; Widowed; Separated; Divorced; Prefer not to answer]

Q38 [**kidsNoAnswer**] Do you have any children aged 18 or under? If so, how old are they? [OPTIONS: No children aged 18 or under [kidsNone]; Yes - children aged under 5 years old [kidsUnder5]; Yes - children aged 5-10 years old [kids5to10]; Yes - children aged 11-15 years old [kids11to15] ; Yes - children aged 16-18 years old [kids16to18]; Prefer not to answer]

Q39 [**householdIncome**] What is the combined annual income of your household, prior to tax being deducted?

[OPTIONS: Up - £7,000; £7,001 - £14,000; £14,001 - £21,000; £21,001 - £28,000; £28,001 - £34,000; £34,001 - £41,000; £41,001 - £48,000; £48,001 - £55,000; £55,001 - £62,000; £62,001 - £69,000; £69,001 - £76,000; £76,001 - £83,000; £83,001 or more; Prefer not to answer]

Q40 [**supportParty**] Generally speaking, do you usually think of yourself as supporter of any one particular party? *[If yes:]*

Q41 [**party**] Which party is that?

[OPTIONS: Conservative; Labour; Liberal Democrat; UKIP; Scottish National Party; Plaid Cymru; Green; Other [Write in]

Q42 [**votedEURef**] Did you vote in the referendum on the UK's membership of the EU held on 23 June 2016? [OPTIONS: Yes; No]

Q43 [**brexitVote**][If Yes:] And how did you vote in the referendum? [OPTIONS: Leave; Remain; Can't Remember]

Q44 [**hypBrexitVote**] [If No:] How do you think you would have voted if you had taken part? Please select one option [OPTIONS: Leave; Remain; Don't Know]

Stage 3 - Undermining message

Experimental design at Stage 4: Randomly split respondents into eight groups: Groups 1 and 2 (control group, 25% of the sample): Proceed directly to stage 4. Groups 3-5 see an undermining message attributed to a professor: Expert bias (3), Personal experience (4), Truthiness (5). Groups 6-8 see an undermining message attributed a a blogger: Expert bias (6), Personal experience (7), Truthiness (8).

Now, we'd like to go back to that earlier statement:

[Re-print fact-checked statement.]

Earlier, you saw some information about that statement. Now, here is a comment on that information from a different source.



David Williams

Professor of Economics at the London School of Economics [*Source: Professor*].

Blogger [*Source: Blogger*].

[*Message: expert bias*] I would take these statistics with a big pinch of salt. The fact that someone is a professor doesn't mean that they don't have an agenda. And we all know that there is a lot of scope to choose and present statistics so that they end up saying just what want them to say.

[*Message: personal experience*] I would take these statistics with a big pinch of salt. A graph might say one thing but the experience of people's everyday lives could be quite different. And I think that a lot of people reading those statistics will say: that doesn't sound like the world I live in.

[*Message: truthiness*] I would take these statistics with a big pinch of salt. There's so much information and so many statistics out there that it can be hard to know what to believe. In that case, I think it's best to trust your instincts even if looks as if the facts are different.

Stage 4

Re-testing belief in false claims

Thinking again about these statements: Would you say they are true or false?

[If respondent saw one of the anti-immigration statements (Q21 or Q22) fact-checked ask them to re-rate the veracity of both anti-immigration claims. If respondent saw one of the pro-immigration statements (Q23 or Q24) fact-checked ask them to re-rate the veracity of both pro-immigration claims.]

Q45 [**noAsylum__T2**] There has been a sharp rise in the number of people applying for asylum in the UK in the past 10 years.

Q46 [**costImmig__T2**] Immigrants receive more in benefits and services than they pay in taxes.

[OPTIONS: 0 Definitely False; 1; 2; 3; 4; 5; 6 Definitely True]

[If respondent saw one of the pro-immigration statements (Q23 or Q24) fact-checked ask them to re-rate the veracity of both pro-immigration claims.]

Q47 [**whiteCrime__T2**] The majority of crimes in London are committed by white people, not ethnic minorities.

Q48 [**lowPaid__T2**] Immigration to the UK does not affect the wages of the low-paid.

[OPTIONS: 0 Definitely False; 1; 2; 3; 4; 5; 6 Definitely True]

Post-truth reasoning

The last two questions asked you to decide whether those statements are true or false. But sometimes it's hard to say whether something is true or false. Sometimes it can be a matter of opinion on which there are different points of view. Thinking again about these statements, and this time on a scale from 0 (purely a matter of fact) to 6 (purely a matter of opinion),

how would you rate the two?

[If respondent saw one of the anti-immigration statements (Q21 or Q22) fact-checked re-print both anti-immigration statements. If respondent saw one of the pro-immigration statements (Q23 or Q24) fact-checked re-print both pro-immigration statements]

Q49 [**noAsylum__opinion**] There has been a sharp rise in the number of people applying for asylum in the UK in the past 10 years.

Q50 [**costImmig__opinion**] European immigrants receive more in benefits and services than they pay in taxes.

[OPTIONS: 0 Purely a matter of fact; 1; 2; 3; 4; 5; 6 Purely a matter of opinion]

[If respondent saw one of the pro-immigration statements (Q23 or Q24) fact-checked]:

Q51 [**whiteCrime__ opinion**] The majority of crimes in London are committed by white people, not ethnic minorities.

Q52 [**costImmig__ opinion**] Immigration to the UK does not affect the wages of the low-paid.

[OPTIONS: 0 Purely a matter of fact; 1; 2; 3; 4; 5; 6 Purely a matter of opinion]

Q53 [**consistent**] Sometimes people see a difference between what they believed was true and what the statistics say. Would you say that the statistics here were consistent with what you believed?

[OPTIONS: Yes, No] *[IF 'Yes', skip to Q43, If 'No' show Q54:]*

Q54 [**ifInconsistent**] Which of these best describes where you stand?

[OPTIONS: The statistics are probably right, but I believe something different; I think that the statistics are wrong; The statistics made me change my mind.] *[randomize order]*

Q55 [**ok2disagree**] And how much would you agree or disagree with the following statement: "It's OK to disagree with the facts if that's what you believe"

[OPTIONS: Strongly Agree [4]; Agree [3]; Disagree [2]; Strongly Disagree [1]]

Trust in the source of the fact-check

Now we would like you to think back about the article you read at the very beginning of this survey, written by Richard Clark, Professor of Population Economics at the University of Oxford (see photo). It provided some information, including a graph about one of the statements you rated on a scale from ‘true’ to ‘false’. *[Re-print Richard Clark’s photo]*

Q56 [**accurate_Expert**] How accurate would you say was the information this professor used?

[OPTIONS: Not at all accurate (1), Not very accurate (2), Fairly accurate (3), Very accurate (4), Can’t remember (5)]

Q57 [**trust_Expert**] And how much would you generally trust what this professor says on the issue of immigration?

[OPTIONS: 0 Would not trust at all; 1; 2; 3; 4; 5; 6 Would trust a great deal; Can’t remember]

[GROUPS 1-2 (who did not see the comment from the 2nd source) end here.]

Trust in the source of the undermining message

[GROUPS 3-5 (who saw a comment from a blogger):]

Turning now to the second source, the comment from David Williams (Blogger, see photo) about Professor Clark’s article:

[Re-print David Williams’ photo]

Q58 [**accuratePostTruthBlogger**] How accurate would you say was the information this blogger used?

[OPTIONS: Not at all accurate; Not very accurate; Fairly accurate; Very accurate; Can’t remember]

Q59 [**trustPostTruthBlogger**] And how much would you generally trust what this blogger says on the issue of immigration?

[OPTIONS: 0 Would not trust at all; 1; 2; 3; 4; 5; 6 Would trust a great deal; Can't remember]

Trust in the source of the undermining message

Groups 6-8 (who saw a comment from a professor):

Turning now to the second source, the comment from David Williams, Professor of Economics at the London School of Economics about Professor Clark's article: [Re-print David Williams' photo]

Q60 [**accuratePostTruthProf**] How accurate would you say was the information this professor used?

[OPTIONS: Not at all accurate; Not very accurate; Fairly accurate; Very accurate; Can't remember]

Q61 [**trustPostTruthProf**] And how much would you generally trust what this professor says on the issue of immigration?

[OPTIONS: 0 Would not trust at all; 1; 2; 3; 4; 5; 6 Would trust a great deal; Can't remember]