Misinformation in the Age of AI

What a decade of fact-checking got wrong, and why AI alone won't fix the problem

PATRICK FLYNN
Senior Data Journalist

FOCALDATA

Contents

Introduction

Last summer, we at Focaldata produced a report on why support for overseas aid had declined over recent decades. In analysing some of the open-text survey responses gathered during the fieldwork process, I was struck by the extent to which people's false perceptions of aid cropped up: where it went, who it went to, and how much was sent.

A decade on from the initial explosion in the debate around misinformation, the problem shows no sign of receding. Many of our clients in the non-profit and media spaces tell us that misinformation remains one of the biggest challenges facing their sector. So how did we get here? Weren't 'fact checkers' supposed to solve everything?

The purpose of this white paper is to understand the misinformation problem as it exists today, beginning with a retrospective on attempts to counter misinformation over the last 10 years, before exploring the drivers of misinformation susceptibility in an evolving media landscape, scoping a typology of misinformation, modelling how different claim types and rhetorical techniques can have outsized negative impacts on vulnerable sectors, and finally looking ahead to how AI is likely to transform the debate over the coming decade.

We set out to test two hypotheses with this research, both of which have been borne out by the findings. The first is a slightly provocative stance, that mainstream attempts to counter misinformation over the last decade – through fact-checking, content labelling, and other methods – have fallen short in their (noble) goals of reducing the amount of misinformation people encounter and their susceptibility to the problem when it does arise. This is covered in the Checked Out chapter of this report. As such, a new paradigm is needed.

The second is that certain kinds of claims are more effective at dealing damage to vulnerable industries like charities and the healthcare sector, particularly 'truth-adjacent' and 'false generalisation' claims which are themselves untrue or unprovable, but draw upon real facts and kernels of truth.

In order to test this second hypothesis, we designed an experiment in which survey participants were shown a variety of claims on a sliding scale of truthfulness – ranging from neutral facts to plainly false claims – and asked to rate their perception of each claim's truthfulness, how a claim

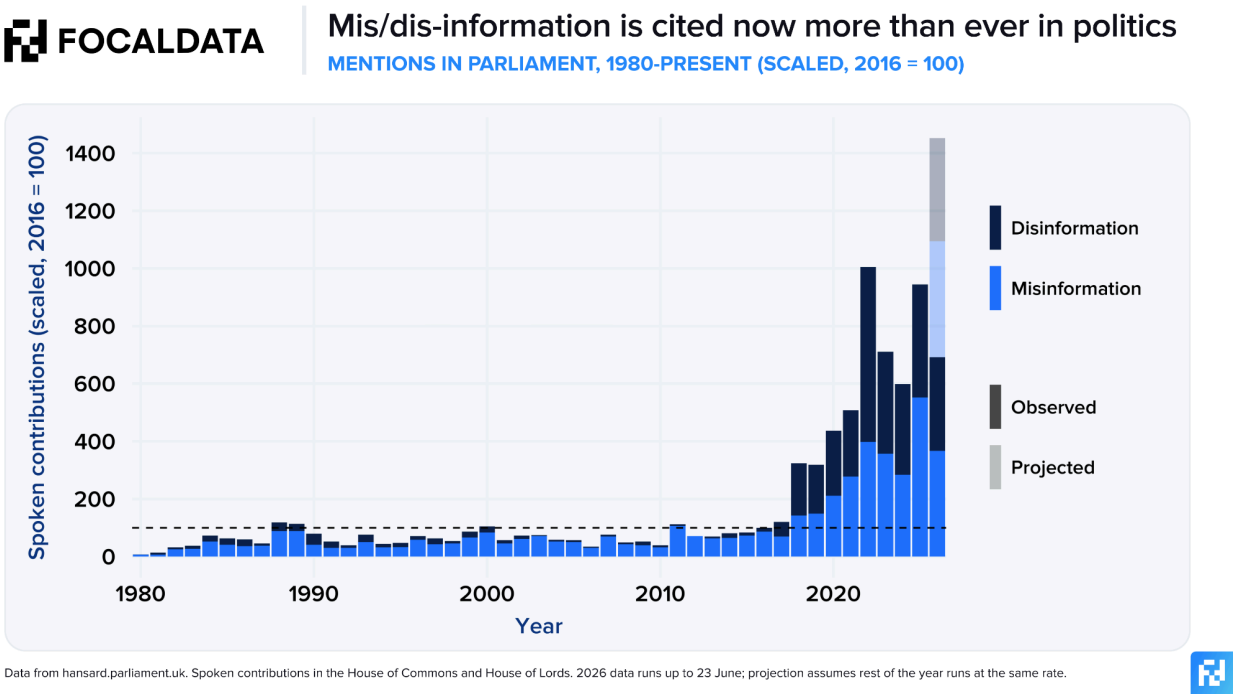
would change their views on a particular sector or group if it were true (impact), and how likely they would be to share it if they believed it to be true (virality). We covered three main problem areas in which misinformation has been particularly prevalent in recent years: charities, healthcare, and migration.

The aim of this report is not to engage in a partisan diatribe about one side of the political spectrum peddling misinformation while the other side bravely fights in pursuit of truth and reality. We tested claims spanning the political spectrum, through which we found that susceptibility to false beliefs was prevalent across the left-right divide.

Artificial intelligence could now pull misinformation in several competing directions at once. The increasing sophistication of AI-generated deepfake images and videos may break the boundaries of reality and fiction entirely, or the growing use of AI chatbots for claim verification and subject research could result in a depolarised and more informed populace. Our data indicate that both of these paths are likely to see the age curve of susceptibility to misinformation reverse, with misinformation becoming a much larger problem for older generations, sparking questions about how the problem is resourced going forward. The government's recent social media ban for teens was driven at least in part by concerns around social media misinformation, but perhaps this is a backward-looking strategy for the late 2010s rather than a solution fit for the 2030s.

The Misinformation Explosion

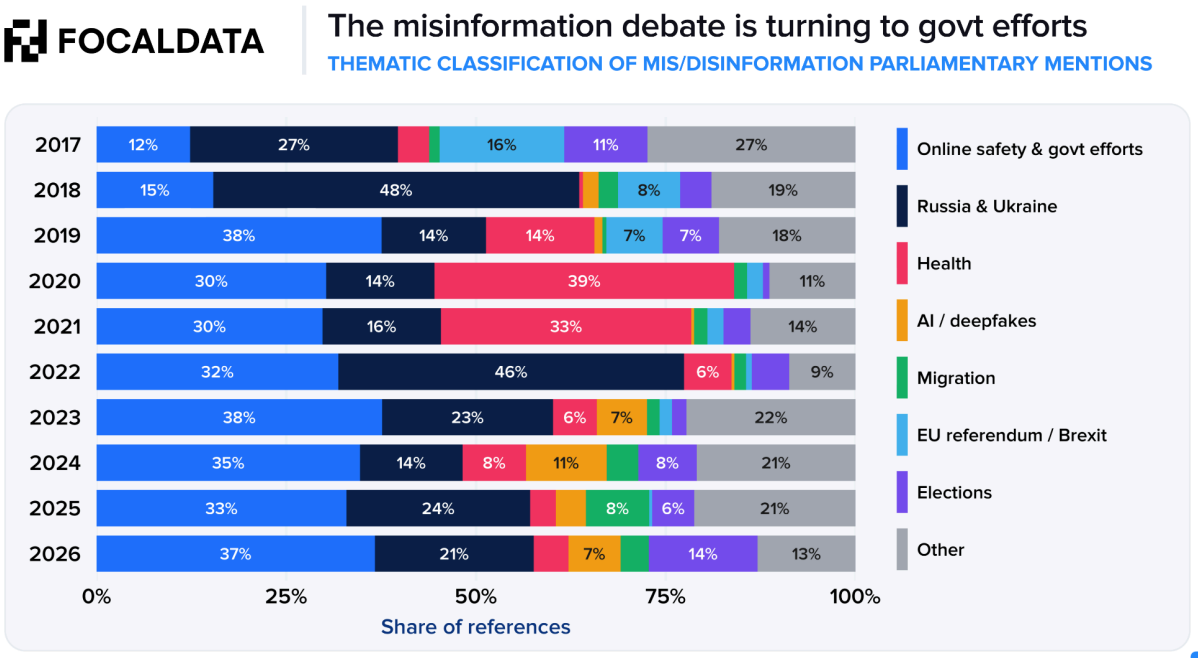
Given how ubiquitous these two terms have become, it is easy to forget that the debate around misinformation and disinformation in public discourse is a relatively new one. The rate at which they were used in the UK Parliament was only as high in 2016 as it was in the late 1980s (just over one mention per week). Since then, the debate has turbo-charged, and if the second half of 2026 continues at the same rate as the first, the record for parliamentary mentions in a given year will be beaten by 40% this year, at a rate 14 times that of 2016.



When separating the two terms, the use of ‘disinformation’ emerges as an almost exclusively post-2016 phenomenon. Since 2022, ‘disinformation’ has been used more times than ‘misinformation’ in Parliament. While some may use the two interchangeably, ‘disinformation’ implies a malicious intent behind its dissemination, something not necessarily applied to its more general synonym. The gradual linguistic shift highlights the growing concern in the political sphere about information warfare and false information being used as a tool for division and undermining democracy.

Tracking the content of these parliamentary contributions over time shows how the debate has moved in the last decade. I used Claude's Haiku model to individually classify thousands of speeches, questions, interventions and statements into thematic groups. Early discussion from 2017–18 focused broadly on Russia and the 2016 European Union referendum, together contributing to a majority of mentions in 2018 (56%). Comments on health misinformation peaked during the COVID-19 pandemic (39% of mentions in 2020; 33% in 2021), and in 2022 Russia and Ukraine contributed to almost half of that year's mentions (46%).

Since then, we have seen increased discussion about migration misinformation following the riots and protests between 2024 and 2025, and AI has made a dent in the numbers after the release of ChatGPT in November 2022.



Data from hansard.parliament.uk. Spoken contributions in the House of Commons and House of Lords. 2026 data runs up to 23 June. Thematic classification by Claude Haiku 4.5.

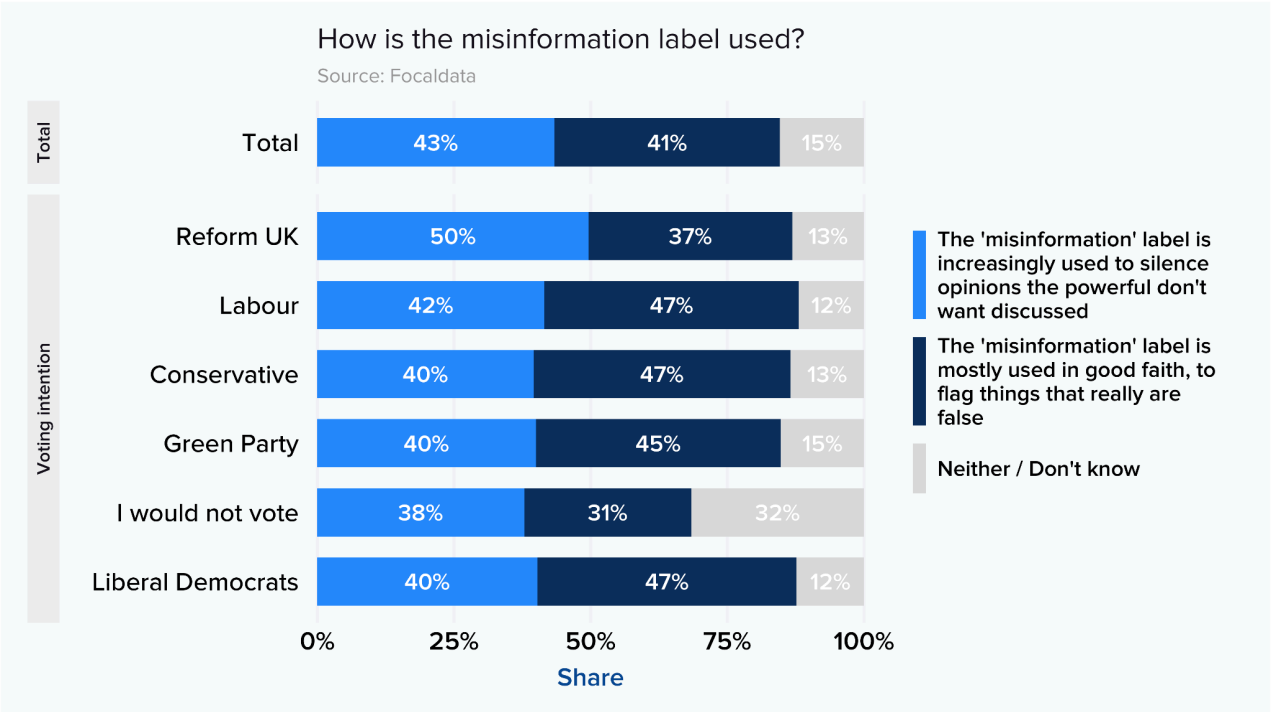
The results of both the above analysis and the survey data to follow suggest that attempts to counter misinformation over the last decade have made little change to the problem. Over the last four years, at least a third of parliamentary references have centred on online safety and the government's attempts to combat misinformation. That the discussion is gradually starting to recentre itself around top-down efforts makes this an opportune time to reflect on the last 10 years and explore options for the future.

Checked Out: Has fact-checking failed?

The public verdict on attempts to counter the misinformation problem over the last 10 years is mixed.

In June 2026, we surveyed 2,326 respondents across Britain, and found that the term ‘misinformation’ has itself become politically contested. The data point to major concerns among the public that the misinformation label is not used in good faith. There should be significant cause for alarm in fact-checking organisations that, for many people, misinformation labelling is perceived as a way to shore up institutional power rather than achieve its purported goal of countering falsehoods.

Rather than asking an agree-disagree question where acquiescence bias could muddy the results, the survey presented respondents with two broadly opposing statements. One suggested that the misinformation label is ‘mostly used in good faith, to flag things that really are false’, while the other posited that the label is ‘increasingly used to silence opinions the powerful don’t want discussed’. The plurality of respondents chose the latter statement, by a narrow margin of 43% to 41%.

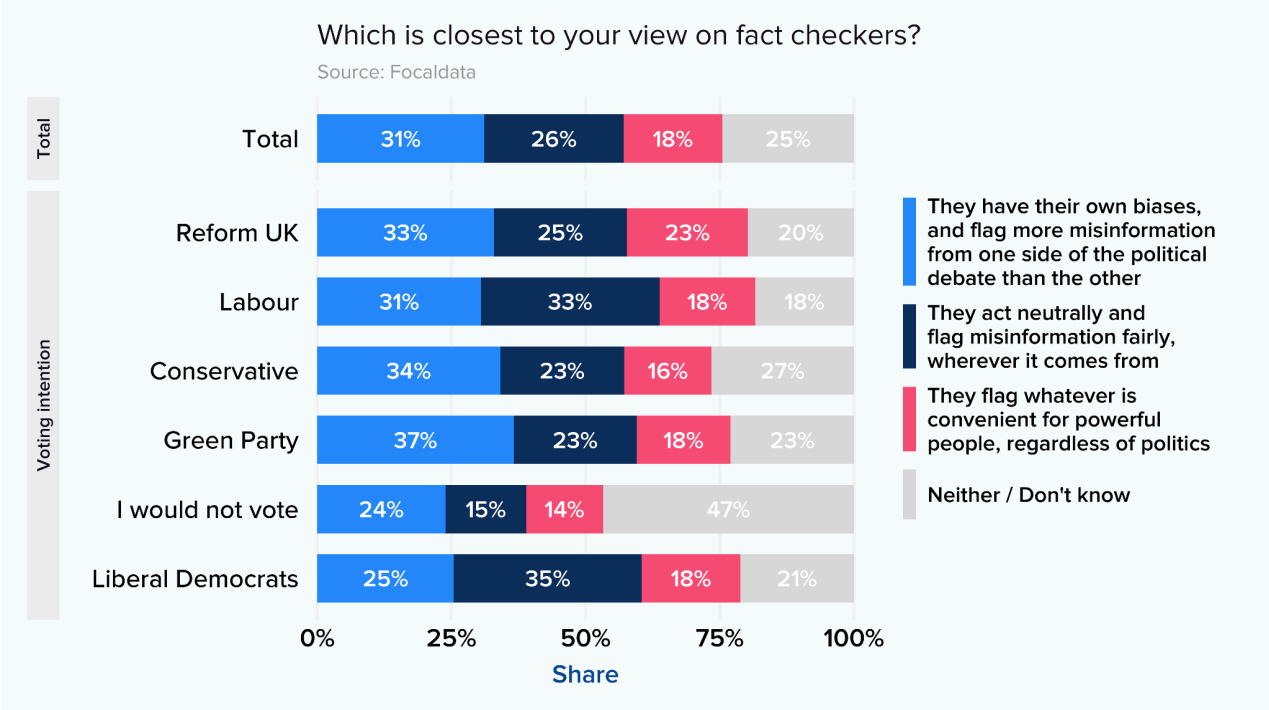


Reform UK voters were the only political group for whom a clear plurality believed in the misinformation-as-silencing-tool claim, but an almost uniform 2-in-5 respondents from all major

parties endorsed that belief. This perception cuts across party lines and therefore cannot easily be fixed by a mere liberal or conservative pivot to counter-misinformation approaches. The public view on misinformation labelling might be better understood as an up-down axis, between the powerful and the powerless, rather than a left-right one.

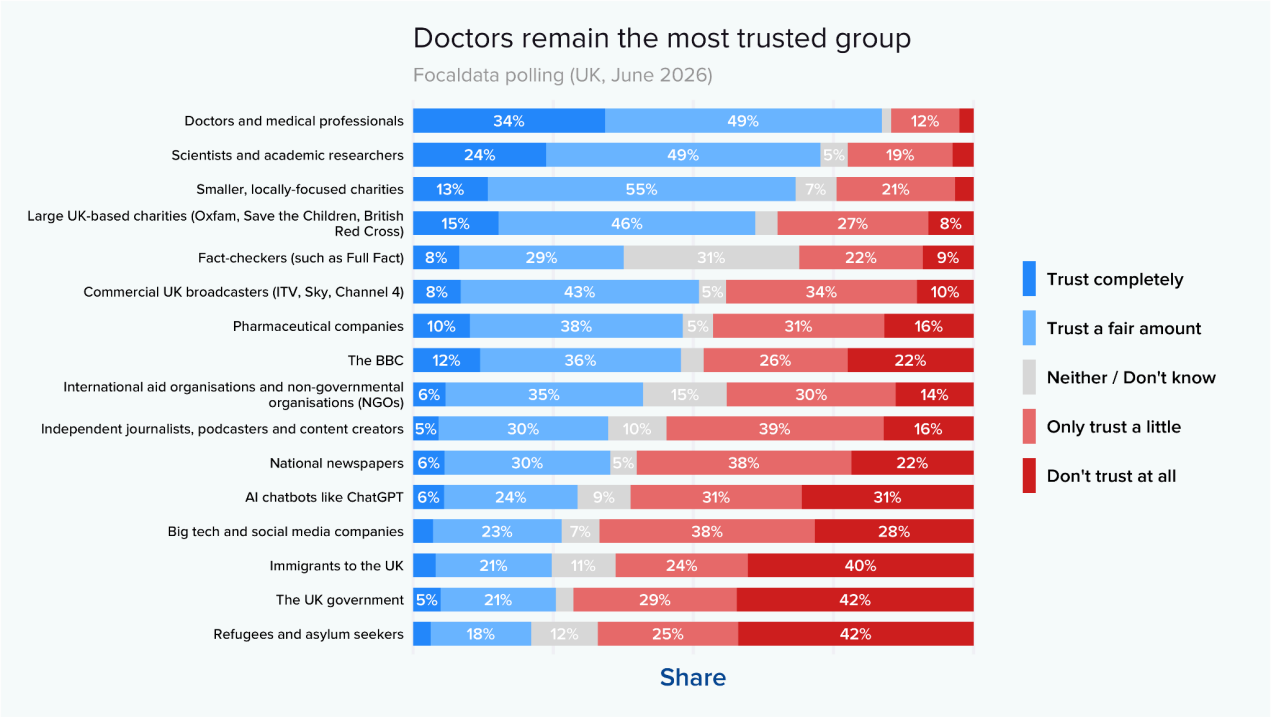
There have been recent reported cases of senior politicians using the term misinformation, or synonyms of it, to try and shut down discussion of inconvenient truths or claims which have later gone on to be proven accurate. It is possible that the public rejection of misinformation labelling is somewhat influenced by these occurrences, but the results on fact-checkers specifically are just as sobering.

Just a quarter of respondents in the survey said they thought fact-checkers act 'neutrally and flag information fairly', while the plurality (31%) believe fact-checkers have their own biases and 'flag more misinformation from one side of the political debate than the other'. The remaining 18% suggested fact-checking biases had no political lean and merely served the interests of those in power, whoever happens to hold it at the time. Once again, these perceptions cut across party lines, with distrust of fact-checkers highest with Reform UK and Green supporters.



Fact-checking organisations can take no comfort in the belief that opposition to misinformation labelling merely reflects a condemnation of politicians who use the term for disingenuous reasons. There is now a large segment of the population, across left and right, which believes that mission creep has occurred: that the category of misinformation has expanded to the point where it is no longer used in good faith, and the definitional drift is now covering what they perceive to be legitimate opinions.

By-and-large, fact-checking organisations have not managed to effectively build trust with the public. Though outright hostility is relatively low, 31% of respondents were either unsure or ambivalent about their trustworthiness – more than twice that of any other group we tested.

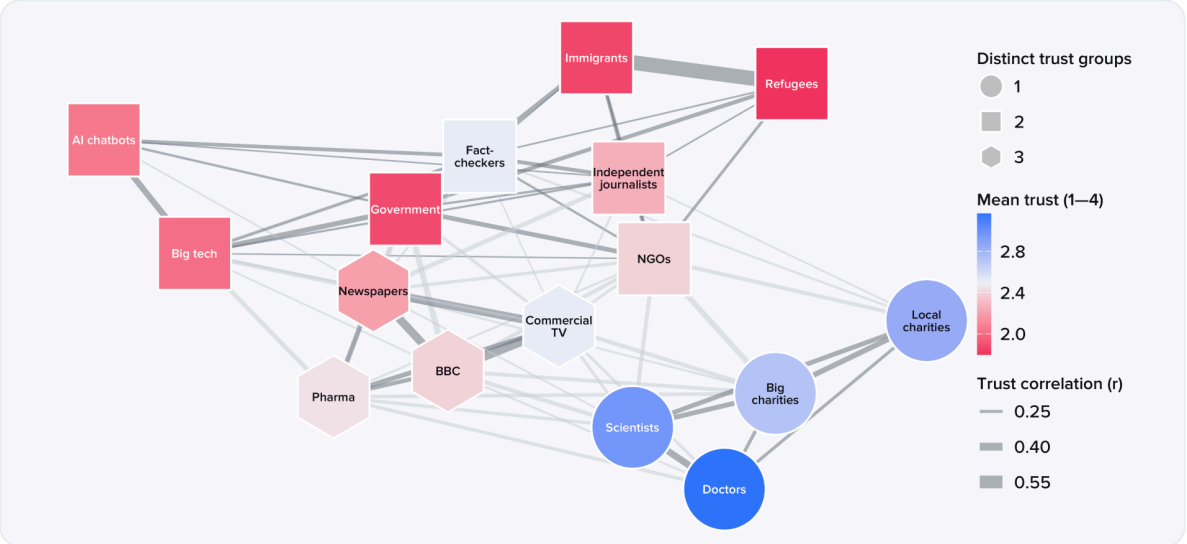


Trust levels in fact-checkers cluster most closely with low-trust and politically-contentious institutions like the government and big tech. The chart below shows network analysis of the correlation between trust levels in various institutions and groups. Nodes are positioned closest to groups where they have the highest correlations. A strong link indicates that the more likely a person is to trust (or distrust) e.g. scientists, the more likely they are to trust e.g. doctors.

FOCALDATA

Fact-checkers cluster with low-trust institutions

NETWORK GRAPH OF INSTITUTIONAL TRUST CORRELATIONS



Fieldwork conducted 4-15 June 2026, with a sample size of 2,326 respondents. Results weighted by age, gender, region, education, ethnicity, recalled 2024 vote and political interest.

While fact-checkers have a mixed overall trust rating and are far from the nadirs of some other groups, the network analysis demonstrates that they remain some distance from the high-trust community containing doctors, scientists, and charities. The current position creates a difficult (perhaps even impossible) environment in which to thrive.

Clearly, fact-checkers will need to play some role in future counterstrategies against misinformation, but perceptions will need to be rebuilt and other avenues pursued. An effective fight against misinformation simply cannot be perceived as a way for the powerful to suppress dissenting views. Winning public trust is hard enough. Gaining enough trust to actively change a person's mind on something they believe is an even more difficult task. The data are clear that this level of trust has not been built over the last decade.

A Typology of (Mis)Information

At the heart of the survey was an experiment designed to test the impacts of different claim types using an original typology influenced by some of the academic literature around misinformation, alongside Focaldata's own work in recent years.

The final typology of information covers nine main categories of claims including neutral facts, critical opinions which cannot be proven true or false, and a set of types ranging from the misleading to the outright false. Some of these terms have a degree of overlap, but each of the 46 claims in the experiment was designed to mostly fall into one of these nine buckets.

This paper is not seeking to label all of these categories as 'misinformation' or 'disinformation'. Aside from the verifiably true and false claims, many of these categories are either unverifiable or contain elements of both truth and falsehood.

Neutral fact

A true, verifiable claim. Uncontested.

Example: "COVID-19 vaccines were approved for use in the UK by the Medicines and Healthcare products Regulatory Agency (MHRA)."

Deceptive implicature

A statement (or statements) which might be factually true or very close to the truth, but is constructed to lead the reader or listener towards a false or unsupported conclusion, for example by exploiting the expectation that facts are presented together because they are linked.

Example: "Last year, while record numbers of British families were forced to use food banks, the chief executives of Britain's twenty largest charities were paid salaries and bonuses totalling more than four million pounds."

Critical opinion

A strongly-held viewpoint rather than a verifiably true or false claim.

Example: “Multiculturalism has failed. Decades of mass immigration have left Britain more divided, with separate communities living parallel lives.”

False generalisation

A real event or anecdote extrapolated to imply a general pattern which is not supported by the example given.

Example: “A refugee who arrived in the UK with nothing went on to qualify as an NHS doctor and save hundreds of lives, which shows that the people coming here are overwhelmingly skilled professionals who quickly give back.”

Truth-adjacent

This technique is similar to deceptive implicature, but differs because the statements are fused together rather than merely presented side-by-side. The very act of fusing or synthesising turns two true statements into a false one, with the bridge between the statements acting as the agent which ‘flips the sign’. ‘Truth-adjacent’ has a double meaning: the claim is both adjacent to the truth, and contains two truths adjacent to one another.

Example: “The UK is sending overseas aid to fund India's space programme.”

In this example, the claim itself is false. It is true, though, that (up until 2015 at least) the UK sent overseas aid to India. It is also true that India has a space programme. Connecting those two statements via the ‘to fund’ bridge turns the statement into a false one. This type of claim is particularly pernicious because it is difficult to counter, and puts the opponent in a no-win situation. Any public figure or organisation who tries to oppose a claim like this would then be asked to justify – rightly or wrongly – why they are defending a situation in which the UK is sending aid to a country with space exploration capabilities.

Motive attribution

A verifiable fact with an unprovable sinister motive bolted on to it.

Example: “Healthcare companies fund vaccination programmes in poorer countries to gain control over their populations.”

In this example, healthcare companies funding vaccination programmes in poorer countries is a verifiable fact and reflects well on the actor involved, but ‘to gain control over their populations’ adds the unprovable implication of a sinister plot behind the action.

Conspiracy theory

Alleges a co-ordinated hidden plan orchestrated by powerful actors. ‘Conspire’ derives from the Latin conspirare (literally: ‘to breathe together’), which gets at the silent or secretive nature of the action. Like motive attribution, it also assigns a sinister motive, but differs because no part of the claim in question is verifiably true.

Example: “Senior figures in the World Health Organization coordinated with pharmaceutical companies to extend the COVID-19 pandemic in order to maximise vaccine sales.”

Imposter source

A fake finding attributed to a real, authoritative source to piggyback off its credibility. An act of stolen valour.

Example: “A leaked Charity Commission internal report has identified more than a dozen major aid charities where senior executives have authorised bonus payments from restricted donor funds.”

Strong falsehood

Flatly false, with no kernel of truth. Verifiably untrue.

Example: “Asylum seekers receive higher weekly payments than those on the state pension.”

Attribute table

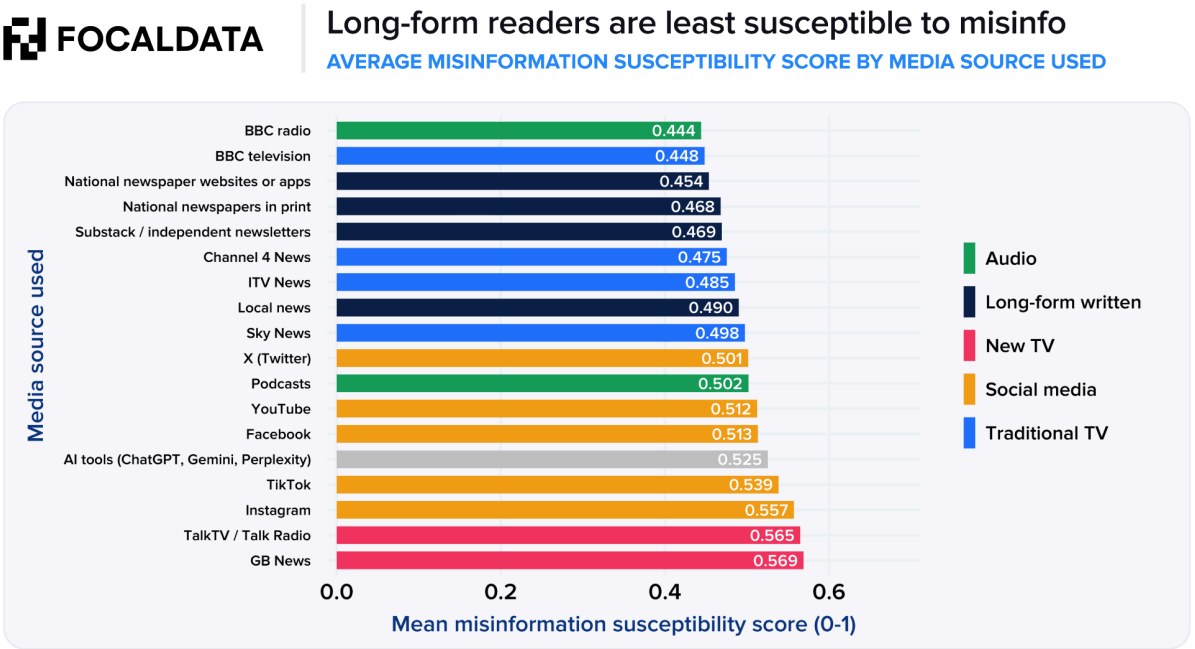
Claim type	Truthfulness	Verifiability	Deceptiveness
Neutral fact	High	High	Low
Deceptive implicature	High	High	High
Critical opinion	N/A	Low	Low
False generalisation	Mixed	Mixed	Low
Truth-adjacent	Mixed	High	High
Motive attribution	N/A	Mixed	Mid
Conspiracy theory	N/A	Low	Mid
Imposter source	Low	High	High
Strong falsehood	Low	High	Mid

What Drives False Beliefs?

Before getting into the results of the experiment, let's first look at the drivers of false or misleading beliefs, which should provide a number of helpful data points for designing effective counterstrategies.

Each respondent in the survey was scored based on their ability to detect facts from falsehoods or unsubstantiated claims, with the most perceptive respondent scores at 0 and the least perceptive at 1. The scoring system was devised to estimate each respondent's misinformation susceptibility levels relative to other survey participants rather than how good they are at detecting false information in general, where an opinion poll can only go so far.

When it comes to media consumption, those who engage with BBC content (radio and television, ranked first and second, respectively) were the most discerning.



Fieldwork conducted 4-15 June 2026, with a sample size of 2,326 respondents. Results weighted by age, gender, region, education, ethnicity, recalled 2024 vote and political interest.
Survey respondents scored based on their ability to separate factual claims from unsubstantiated ones, where 0 = most discerning respondent and 1 = least discerning respondent.

A broad trend in the data emerged, where the older or more traditional a media type is, the less susceptible its users are to misinformation. Though individual sources vary in ranking (BBC television viewers score better than other traditional TV channels), the average ranking by category is effectively a timeline of media: long-form written content (Substack; newspapers) → audio (radio;

podcasts) → traditional television (BBC; ITV News; Channel 4 News) → social media (Facebook; TikTok) → new, 'Americanised' TV news (GB News; TalkTV).

Of the nine media sources whose users were more susceptible than the average respondent, seven of these were social media platforms or new television news channels. There was no social media platform whose users were more discerning than an average respondent.

The results provide more evidence that misinformation susceptibility, short-form media usage, populist beliefs, and shorter attention spans are all children of a shared cultural nexus arising in the late 2010s through to the 2020s in Western culture. The Medium is the Message.

AI may take this era in a new direction, but there is currently little evidence that AI tools are bucking the trend of newer technologies having more misinformation-liable users. Users of AI chatbot tools are currently no better than social media users at detecting unsubstantiated claims.

One of the striking things about the media consumption finding is that it holds up even after accounting for a variety of demographic variables which covary with media habits, including age, gender, and education levels.

In addition, the pattern holds after directly controlling for critical-thinking abilities. The survey presented respondents with three logic questions:

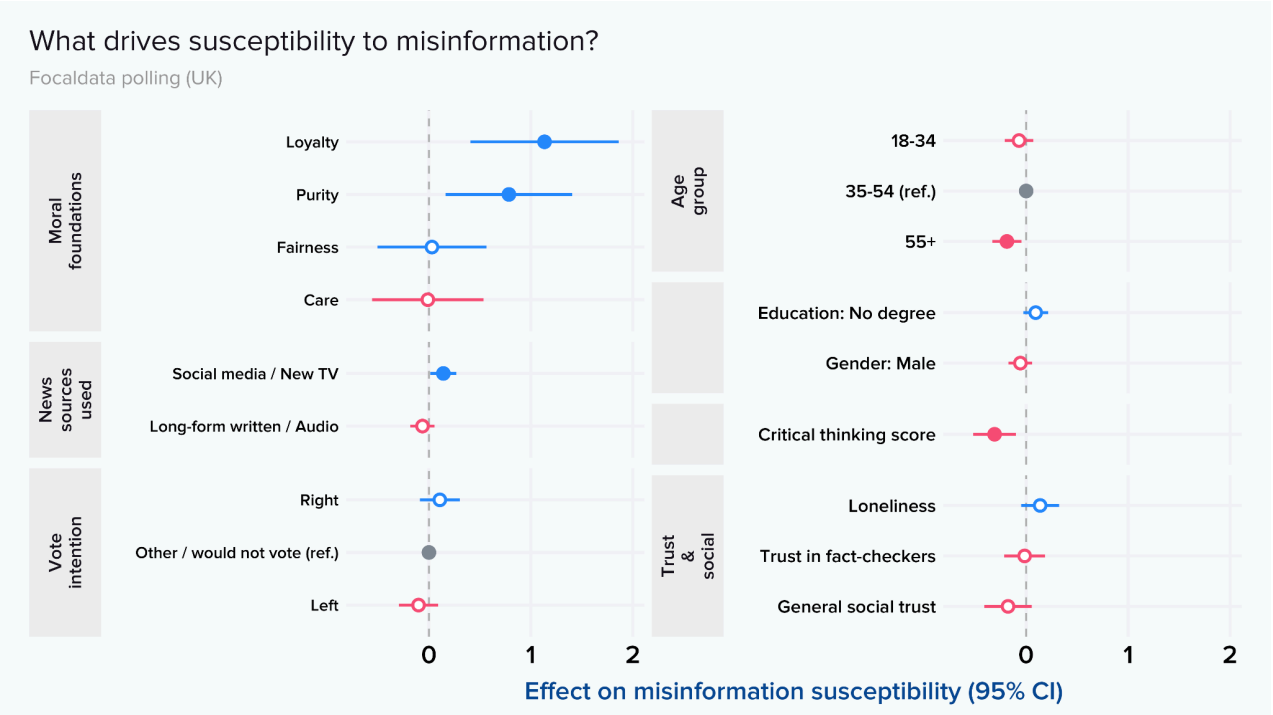
- "A cricket bat and a ball together cost £1.10. The bat costs £1 more than the ball. How much does the ball cost?"
- "A patch of lily pads doubles in size every day. If it takes 48 days to cover the whole lake, how long does it take to cover half the lake?"
- "Imagine you are running a race. If you overtake the person in second place, what position are you now in?"

Each respondent was assigned a percentage score based on how many of the questions they answered correctly (5p, 47 days, and second place, though I am sure the reader knew that already). Just 5% of panellists correctly answered all three questions, and a further 8% got two correct answers.

Those who answered more of the logic questions correctly were significantly more discerning when it came to separating neutral facts from unsubstantiated claims. Both critical thinking abilities and

media sources used remain significant in a larger model predicting an individual's susceptibility to false information.

A 30-year-old, degree-educated, female respondent who only reads newspapers and possesses high critical-thinking skills is significantly more likely to be able to detect misinformation than a 30-year-old, degree-educated, female respondent who possesses high critical-thinking skills and gets her news from Instagram.



In addition to news sources and critical thinking skills, the large effects of moral foundations stand out in the model, specifically loyalty and purity. Respondents who were more likely to say whether or not someone 'showed a lack of loyalty' or 'violated standard of purity and decency' were important drivers in how they decide whether an act is right or wrong were significantly less likely to be able to separate facts from unsubstantiated claims.

Similar to the potential connection between long-form media and education levels, moral foundations have a strong overlap with political identity. Once again, though, the former driver holds up even after controlling for the latter. In fact, there was no significant effect of voting intention or political belief on misinformation susceptibility when other variables have been accounted for.

While moral foundation beliefs affect both voting intention and susceptibility to misinformation, it is those moral foundations, rather than voting intention, that appear to spur susceptibility.

It may be the case that the effects of moral foundations in the model can be partially explained by the subjects of the claims we tested (charities, healthcare, and migration). As Jonathan Haidt, one of the creators of the moral foundations theory, states on the moral foundations website, the purity foundation was designed around the 'psychology of disgust and contamination', which lies behind ideas of the body as a 'temple which can be desecrated by [...] contaminants'. It stands to reason that people who place relatively high levels of value on contamination and desecration in their moral understanding of the world would be more likely to believe unsubstantiated claims around healthcare – particularly vaccines, which typically involve injecting weakened forms of a disease-causing substance into the body. A third of the survey's claims centring on healthcare could therefore artificially inflate the effects of the purity moral foundation in the model. Of the data collected, however, there was no significant relationship between the purity foundation and levels of misinformation susceptibility by topic.

Moral foundations, then, ought to be considered more thoroughly in the misinformation debate, on top of traditional demographics or even where people get their information. These underlying beliefs can strongly shape an individual's susceptibility to misinformation, apparently even more so than political ideology, and should be addressed directly in future counterstrategies.

The Big Leap: Experiment results

The primary focus of the experiment was to determine the **current impact** of each of the 46 claims (how much it had changed public opinion on a particular sector among those who had heard it before and believed it to be true), the **potential impact** (how much it would change views if everyone had heard it), and the **potential viral impact** (potential impact multiplied by a virality score – i.e. how likely people would be to share it if they believed it were true).

Impact scores were calculated using the following formulae:

*Current impact: Ih * Th * H*

*Potential impact: I * T*

*Potential viral impact: (I * T) * V*

Where:

***Heard before (H)** = Share of respondents who had heard the claim before*

***Maximum impact (I)** = Overall favourability change towards [migrants / healthcare industry / charities] if claim was true and every respondent had heard the claim*

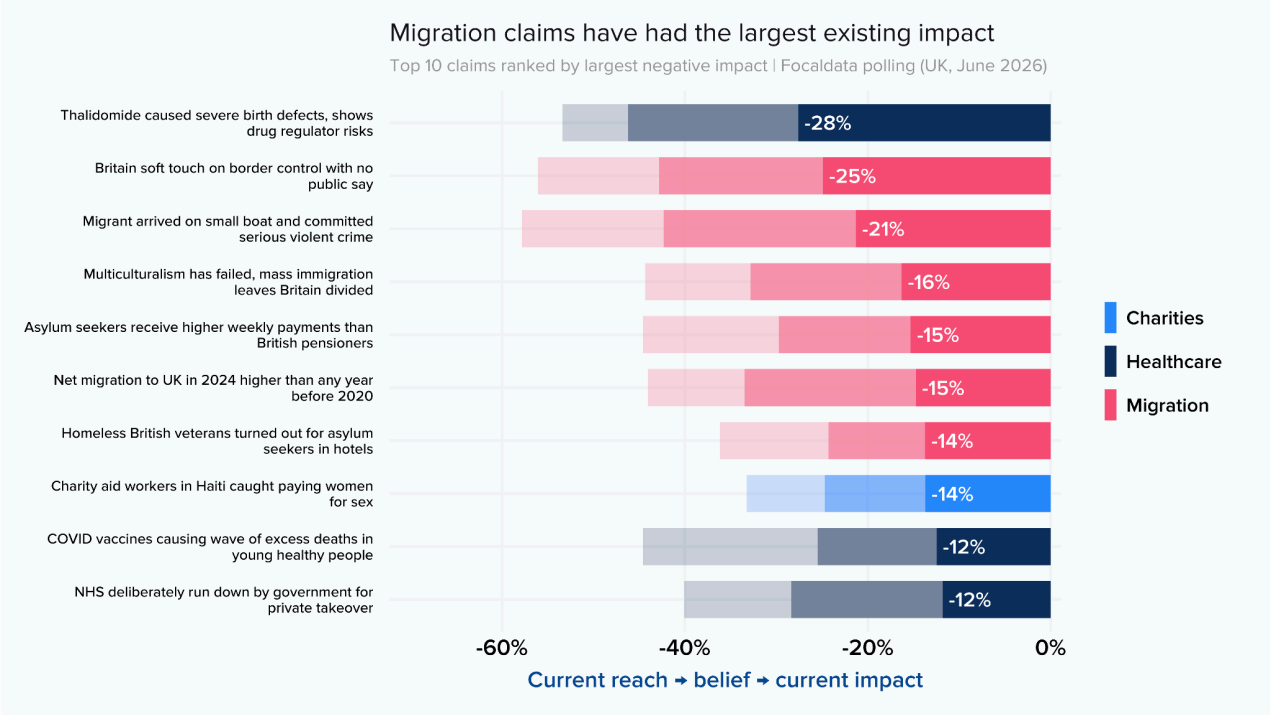
***Impact heard (Ih)** = Favourability change towards [migrants / healthcare industry / charities] if claim was true, among those who had heard the claim before only*

***Likely true (T)** = Average truth likelihood assigned to each claim (between 0% and 100%)*

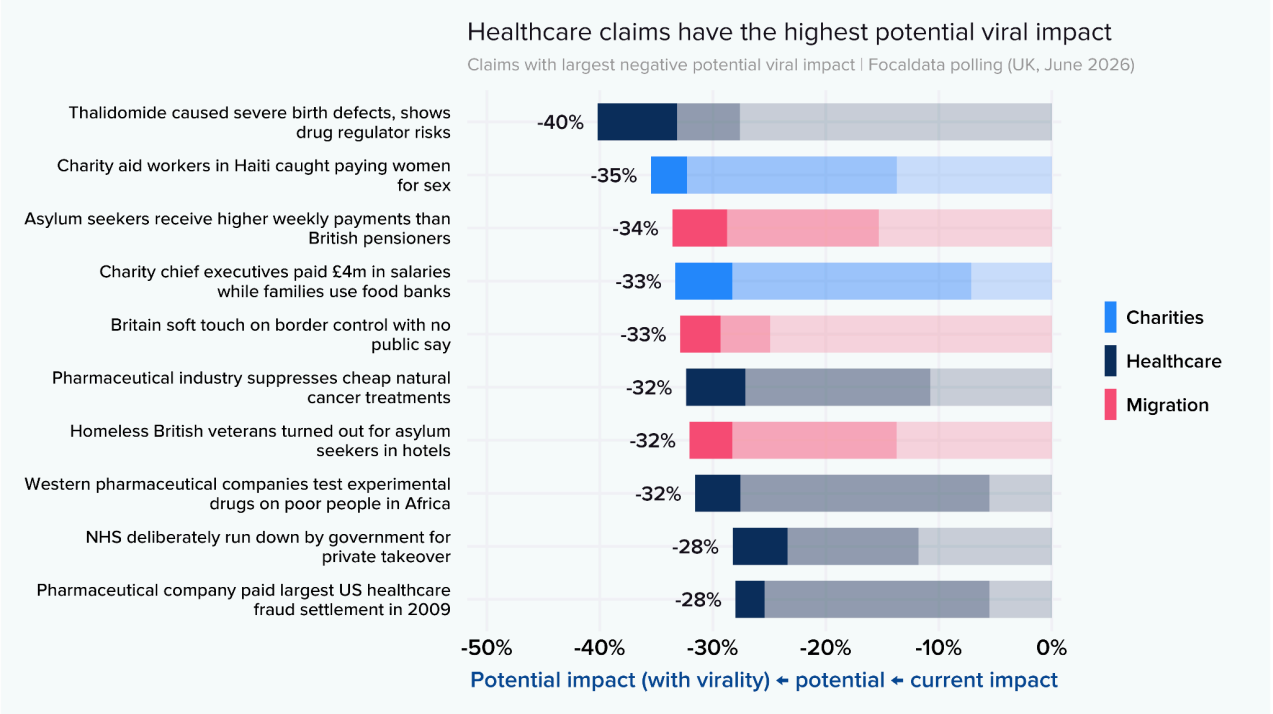
***Likely true heard (Th)** = Average truth likelihood assigned to each claim, among those who had heard the claim before only*

***Virality (V)** = Multiplicative average likelihood of a respondent sharing a claim if they believed it to be true, where 2 = Very likely and 0 = Not at all likely.*

Across the 46 claims we tested, those covering migration have had by far the largest current negative impact on views of migrants to the UK (versus views on the charity sector or healthcare industry). Of the 10 claims with the highest current impact, six focused on migration. To reiterate, not all of the claims we tested are outright misinformation – some are factual, some are opinions-based, and others contain elements of truth.



When we look at **potential viral impact**, healthcare dominates with five of the top ten claims with the highest scores. Claims around migration seem to be baked into public opinion already, whereas those concerning healthcare carry far more latent risk for the industry, even after the COVID-19 pandemic. Information about Pfizer's 2009 multi-billion dollar fine in the United States and a suggestion that Western pharmaceutical companies routinely test experimental drugs on people in low-income countries both saw muted current impacts, but surged into the top 10 in a modelled scenario where everyone had heard them.



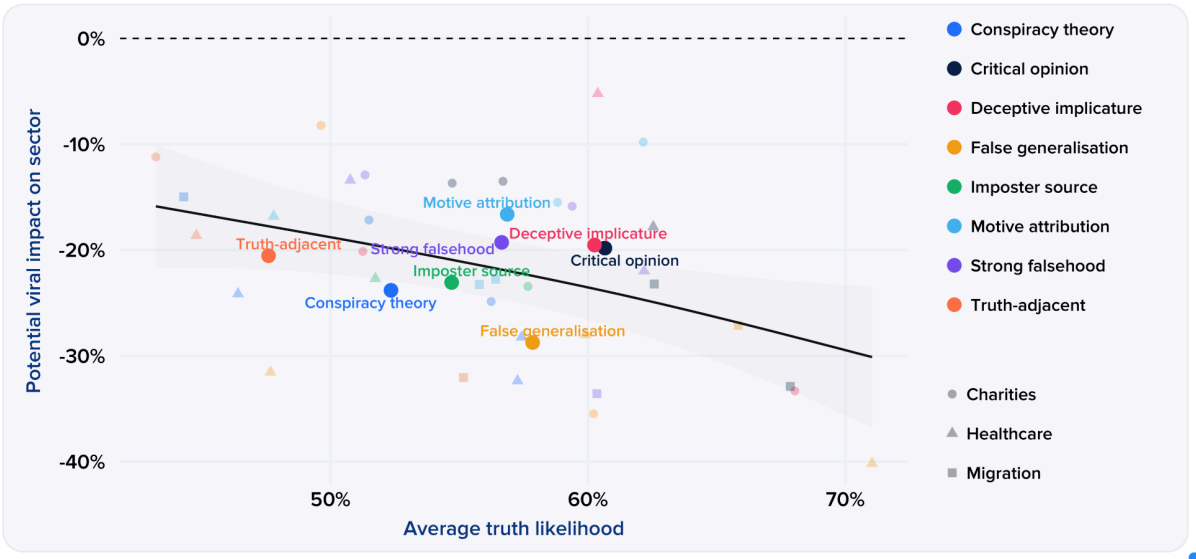
Perceived truthfulness of a claim is the biggest driver of its potential viral impact, so the impact of individual claim types from the typology of (mis)information proposed earlier is somewhat mediated by the specific claims tested. In order to reliably say a particular type of (mis)information is more damaging than any other in and of itself, we would typically need to test hundreds of claims.

To mitigate this, the effects of each claim can be estimated conditional on its perceived truthfulness. In doing so, a few major claim types emerge as overperformers when controlling for believability levels. For the purposes of estimating negative impacts, neutral facts and positively-framed claims have been removed from the chart below.

FOCALDATA

False generalisations appear to do most damage

POTENTIAL VIRAL IMPACTS BY CLAIM TYPE AND SECTOR



Fieldwork conducted 4-15 June 2026, with a sample size of 2,326 respondents. Results weighted by age, gender, region, education, ethnicity, recalled 2024 vote and political interest.

False generalisations have both the largest overall potential viral impact, as well as the largest negative impact relative to perceived truthfulness. That data suggest that it might therefore be anecdotal true events, extrapolated to suggest a wider unproven pattern, which have a more damaging impact to each sector than explicitly false claims.

As seen in the top 10 claims with the largest potential viral impact, the thalidomide scandal from the 1950s and 60s – used to suggest the public should never trust drug regulators – and the Oxfam Haiti scandal revealed in 2018 – used to argue the charity sector is routinely covering up exploitation issues – were ranked as the top two. Both of these claims fall under the ‘false generalisation’ archetype.

Conspiracy theories and truth-adjacent claims both also overperform relative to estimated truthfulness. In the latter group, the suggestion that homeless military veterans are being turned out of hotel rooms so the government can house people arriving on small boats was the most potentially damaging.

Across many of the most impactful claim types, it appears that those which contain explicit dis-/mis-information gain less traction than those which contain kernels of truth. These are the claims that spread most and do the most damage. In the example just given, many people may not

literally believe that local authorities are directly removing homeless people from hotels to make way for asylum seekers, but the belief taps into a perception of injustice in the eyes of the believer. The truth-adjacency is what strikes a chord. Though the claim itself is false, it is true that homeless veterans do not get the same form of housing support, alighting on a 'wider truth' for those who endorse the claim. Like the 'Indian space programme' aid example outlined earlier, claims like these are harder to debunk because the underlying concern it is based upon is not entirely false.

False generalisations, truth-adjacent claims, and conspiracy theories often have the same thing at their heart. Like a piece of popcorn which starts as a kernel and turns itself inside out, a kernel of truth becomes warped into a false claim.

In 'The Paranoid Style in American Politics' from 1964, Richard Hofstadter talked about the 'big leap from the undeniable to the unbelievable'. When considering some of the most popular conspiracy theories throughout history, many have these truth kernels. It is no surprise when understood through this lens that the rise in UFO conspiracy theories in the United States during the 1940s occurred contemporaneously with the Cold War, with widespread cultural anxieties about an alien culture with potentially advanced technologies and the ability to brainwash American citizens. An element of truth or reality becomes warped into something unprovable: from the undeniable, to the unbelievable.

Whose Role Is It, Anyway?

In conversations with people across the media sector ahead of writing this report, many raised concerns about where the responsibility for combating false or misleading information sits. Some argued that the government itself was not entirely clear on their position on this question, even after a decade of parliamentary debate. Others argued that it was the responsibility of social media companies to regulate the content produced on their platforms as if it were their own, and that governments should mandate this.

The vast majority of the public think companies have at least some responsibility for flagging the content that appears on their platform in principle. Only 14% of respondents think false or misleading content should be 'left alone' as free speech should take precedence.

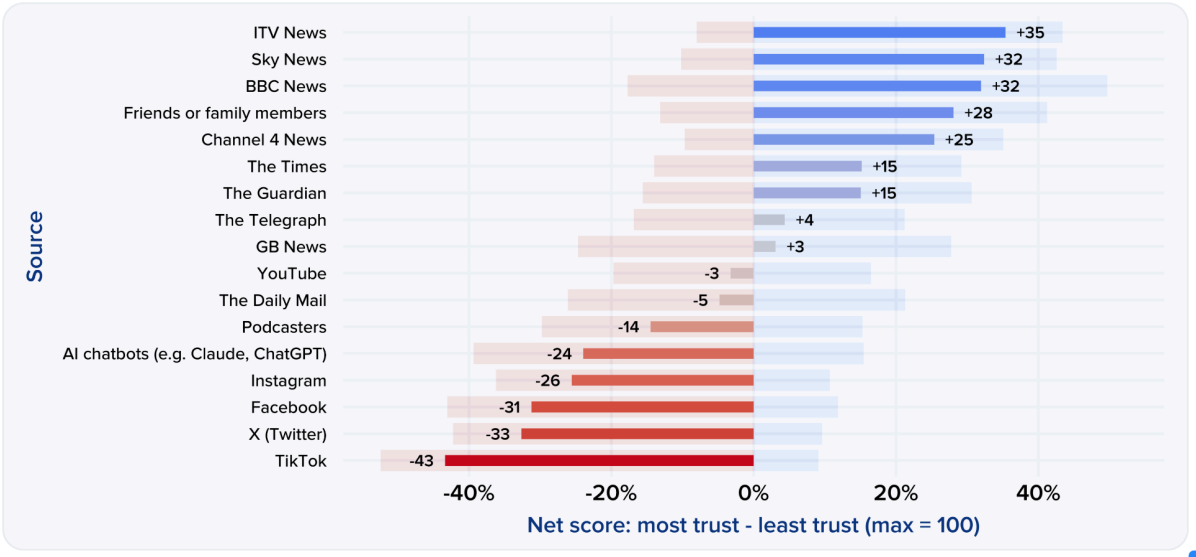
By the time unsubstantiated claims reach the attention of mainstream broadcasters and fact-checkers, those same claims will have already been shared thousands of times across personal networks on messaging platforms and social media sites.

If the responsibility does fall on mainstream news sources, they at least enjoy a fairly large lead in public trust to provide accurate information, with ITV News taking the top spot in our MaxDiff results.

FOCALDATA

ITV is narrowly the most trusted source for accurate info

MAXDIFF RESULTS ON TRUST TO GIVE ACCURATE INFO ON POLITICS/SOCIETY



Fieldwork conducted 4-15 June 2026, with a sample size of 2,326 respondents. Results weighted by age, gender, region, education, ethnicity, recalled 2024 vote and political interest.

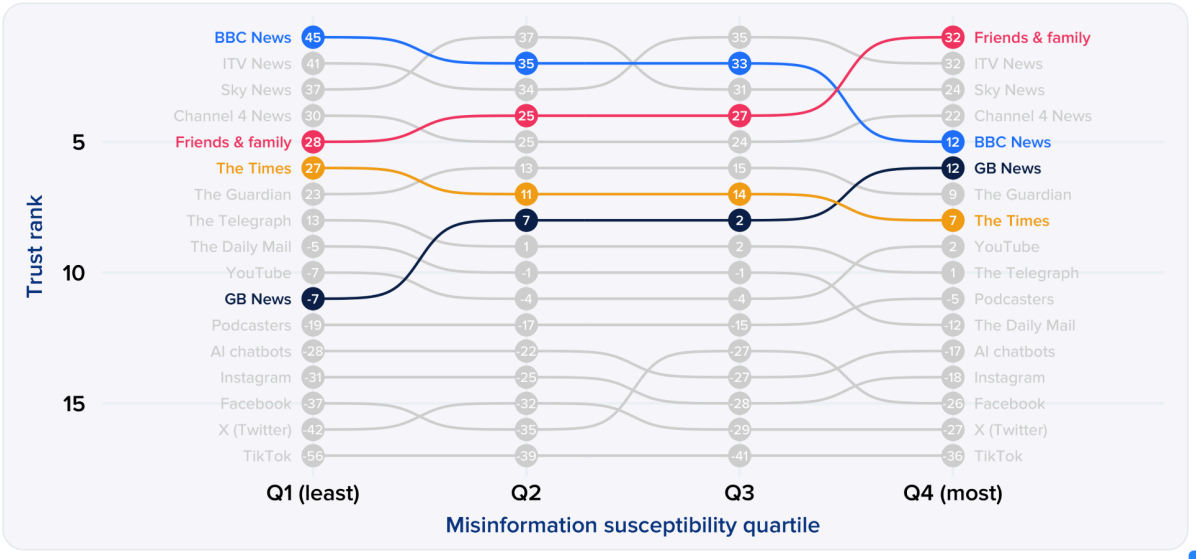
However, some mainstream institutions need to grapple with the fact that their trust levels are lowest with those most susceptible to misinformation. Friends and family, the only non-media source we asked about, rises from fifth place in the most discerning group (17 points behind the leader), to first place with the most susceptible group (tied with ITV News).

Of the major broadcasters, the BBC in particular sees a drop not just in its trust levels, but in its relative scores compared to other news sources, dropping from first place to fifth. This is not a case of the most susceptible groups simply trusting all news outlets less. ITV and Sky both see negative gradients as we move up the susceptibility ladder, but the BBC stands out for a particularly steep drop in trust.

FOCALDATA

BBC trust is lowest with those susceptible to misinfo

MAXDIFF TRUST RANKS & SCORES BY SUSCEPTIBILITY QUARTILE



Fieldwork conducted 4-15 June 2026, with a sample size of 2,326 respondents. Results weighted by age, gender, region, education, ethnicity, recalled 2024 vote and political interest.

The lower levels of trust in established media institutions from the groups most likely to struggle in separating facts from unsubstantiated claims goes some way in explaining why the plurality of Britons think misinformation labelling is being used as a tool by the powerful to suppress information, and why fact-checkers have failed to cut through.

The conversation, then, may have to turn to a question of regulation rather than countering misinformation through existing tools – though, again, this is not to suggest that fact-checking and content labelling should play no role.

A major risk is that governmental attempts to regulate are already behind the curve and are chasing yesterday's problem. Artificial intelligence is not just going to transform how we think about misinformation in the next decade, the transformation is already here.

AI and the Future of Misinformation

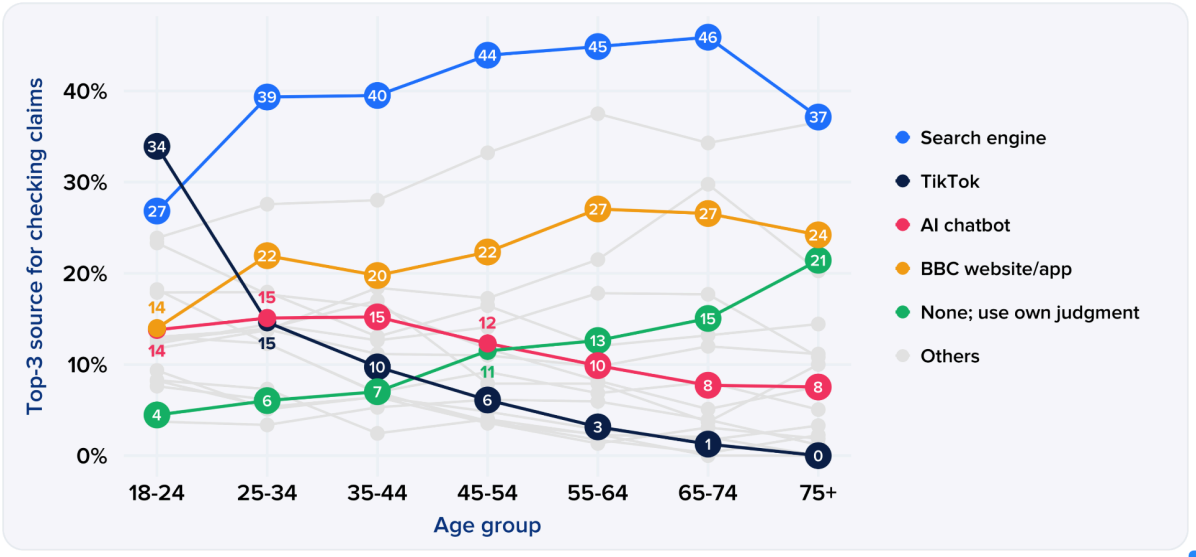
Artificial intelligence could take the misinformation debate in several different directions. On one hand, public trust in the content they see could collapse entirely as the boundaries between real and AI-generated content collapse – this is the AI-negative platform. On the other, LLMs and AI integration in search engines could begin to repair misinformation damage as trusted fact-checkers, with users actively searching for claims and being met with well-researched, balanced results which both inform and depolarise. The FT's John Burn-Murdoch covered this AI-positive pathway in a column back in March.

The more utopian of these two pathways appears to be a long way off, however. Our survey data shows that AI chatbots are rarely used for verifying claims, and few people fully trust their outputs.

FOCALDATA

AI chatbots are rarely used to check claims

SHARE USING EACH SOURCE TO VERIFY CLAIMS (TOP 3), BY AGE GROUP



Fieldwork conducted 4-15 June 2026, with a sample size of 2,326 respondents. Results weighted by age, gender, region, education, ethnicity, recalled 2024 vote and political interest.

Only 12% of respondents currently use LLMs as one of their leading tools for checking claims they are not sure are true or false. Search engines still dominate with all age groups, with the exception of under 25s, where TikTok records a share of 34%, more than double its score with any other age group, and seven points ahead of search engines.

The tech habits of young people often give us an indication of where the rest of society is likely to follow. The data show that the most used verification sources for young people fall outside the same Ofcom rules and regulations that govern mainstream broadcasters. There is an age gradient for LLM usage, but uptake is likely to be slower given that the age curve only peaks at 15% with 25-34 year olds. This 15% should be situated in the context of ChatGPT now being the fifth most visited website in the world. While LLMs are increasingly used, few are engaging with them for claim-verification purposes. However, it is possible that with the greater integration of AI tools into search engines, people will see more AI-generated summaries indirectly, which could have a similar informational effect.

There is stronger evidence that the AI-negative pathway is emerging more clearly. Usage of AI chatbots for countering misinformation falls well below respondents' ability to detect deepfakes. 18-24 year olds are the only age group where a majority say they can usually detect whether a video is real or AI-generated (51%). A plurality or majority of every other age group say they find it difficult to know, climbing to 61% with 65-74 year olds, and all the way to 82% with those aged 75 or older. The rapid progression of AI image and video capabilities is racing far ahead of AI's usage for claim verification.

Regardless of which one emerges most strongly, both the AI-negative and AI-positive pathways are likely to lead to misinformation becoming a larger problem for older age groups. Older Britons are the least likely to use LLMs for verification they see, are the most likely to rely on their own judgment and not verify claims at all, and are the least likely to be able to detect AI-generated video content. The next 10 years of the misinformation debate will need to take this into account.

Final Thoughts

Misinformation 10 years ago was a problem many thought, perhaps naively, could be solved with better fact-checking and clearer content labelling. A decade of experimentation is over, and the public verdict is clear that those tools have not worked. A plurality of Britons now believe that misinformation labelling is a weapon of the powerful, a view which cuts across the political spectrum.

We tested nearly 50 claims for this white paper, categorised under a new typology of (mis)information, and found that it was not outright lies that did the most damage to vulnerable sectors. Instead, it was those which start from a real example and are extrapolated in a number of ways to support an unsupported conclusion. A kernel of truth acts as a catalyst for damaging claims to spread.

Our analysis found that media diet, moral foundations, and critical-thinking scores were strong predictors of a person's ability to separate fact from fiction. Political affiliation was not statistically significant. This finding, alongside the broad political makeup of those who distrust fact-checkers, suggest that aiming to solve for left and right in this debate is not the right choice.

Early evidence suggests that artificial intelligence will not resolve the misinformation crisis any time soon, and may even exacerbate it. Whichever pathway emerges strongest: widespread use of LLMs to verify claims or deepfakes which erode perceptions of visual reality, older generations are likely to be the flashpoint for misinformation susceptibility over the next decade. This group is the least likely to check claims (using LLMs or otherwise) and the least likely to be able to detect AI-generated video content. The government's recent social media ban may be a solution for the previous decade's dynamics than the next one's.

This report does not argue for entirely abandoning fact-checking, but the evidence gathered should be a warning sign for how attempts to counter misinformation can go wrong. Future counterstrategies should make use of the findings that institutional trust needs to be rebuilt with vulnerable communities, and claims with elements of truth do the most damage. Rightly or wrongly, Britons have the perception that the misinformation category has been broadened beyond what is sensible or reasonable.

Ten years on from the explosion in this debate, the problem has not yet been solved, and the ground is already shifting again. Hopefully this report can serve as a useful roadmap for making the next 10 years more successful.

FOCALDATA

FOCALDATA.COM