



ARTICLE



<https://doi.org/10.1057/s41599-024-03503-6>

OPEN

Liars know they are lying: differentiating disinformation from disagreement

Stephan Lewandowsky ^{1,2✉}, Ullrich K. H. Ecker ³, John Cook⁴, Sander van der Linden⁵, Jon Roozenbeek⁶, Naomi Oreskes⁷ & Lee C. McIntyre⁸

Mis- and disinformation pose substantial societal challenges, and have thus become the focus of a substantive field of research. However, the field of misinformation research has recently come under scrutiny on two fronts. First, a political response has emerged, claiming that misinformation research aims to censor conservative voices. Second, some scholars have questioned the utility of misinformation research altogether, arguing that misinformation is not sufficiently identifiable or widespread to warrant much concern or action. Here, we rebut these claims. We contend that the spread of misinformation—and in particular willful disinformation—is demonstrably harmful to public health, evidence-informed policymaking, and democratic processes. We also show that disinformation and outright lies can often be identified and differ from good-faith political contestation. We conclude by showing how misinformation and disinformation can be at least partially mitigated using a variety of empirically validated, rights-preserving methods that do not involve censorship.

¹University of Bristol, Bristol, UK. ²University of Potsdam, Potsdam, Germany. ³University of Western Australia, Crawley, WA, Australia. ⁴University of Melbourne, Melbourne, VIC, Australia. ⁵University of Cambridge, Cambridge, UK. ⁶Kings College London, London, UK. ⁷Harvard University, Cambridge, UK. ⁸Boston University, Boston, MA, USA. ✉email: stephan.lewandowsky@bristol.ac.uk

“If everybody always lies to you, the consequence is not that you believe the lies, but rather that nobody believes anything any longer.”

— Hannah Arendt

One of the normative goods on which democracy relies is accountable representation through fair elections (Tenove, 2020). This good is at risk when public perception of the integrity of elections is significantly distorted by false or misleading information (H. Farrell and Schneier, 2018). The two most recent presidential elections in the U.S. were accompanied by a plethora of false or misleading information, which grew from false information about voting procedures in 2016 (Stapleton, 2016) to the “big lie” that the 2020 election was stolen from Donald Trump, which he and his allies have baselessly and ceaselessly repeated (Henricksen and Betz, 2023; Jacobson, 2023). Misleading or false information has always been part and parcel of political debate (Lewandowsky et al., 2017), and the public arguably accepts a certain amount of dishonesty from politicians (e.g., McGraw, 1998; Swire-Thompson et al., 2020). However, Trump’s big lie differs from conventional, often accidentally disseminated, *misinformation* by being a deliberate attempt to *disinform* the public.

Scholars tend to think of disinformation as a type of misinformation and technically that is true: intentional falsehoods are but one subset of falsehoods (Lewandowsky et al., 2013) and intentionality does not affect how people’s cognitive apparatus processes the information (e.g., L. K. Fazio et al., 2015). But given the real-world risks that disinformation poses for democracy (Lewandowsky et al., 2023), we think it is important to be clear at the outset whether we are dealing with a mistake versus a lie.

The tobacco industry’s 50-year-long campaign of disinformation about the health risks from smoking is a classic case of deliberate deception and has been recognized as such by the U.S. Federal Courts (Smith et al., 2011, see also Civil Action 99-2496(GK) United States District Court, District of Columbia. *United States v. Philip Morris Inc.*). This article focuses primarily on the nature of disinformation and how it can be identified, and places it into the contemporary societal context. Wherever we make a broader point about the prevalence of false information, its identifiability or its effects, we use the term *misinformation* to indicate that intentionality is secondary or unknown.

An analysis of mis- and disinformation cannot be complete without also considering the role of the audience, in particular when people share information with others, where the distinction between mis- and disinformation becomes more fluid. In most instances, when people share information, they do so based on the justifiable default expectation that it is true (Grice, 1975). However, occasionally people also share information that they know to be false, a phenomenon known as “participatory propaganda” (e.g., Lewandowsky, 2022; Wanless and Berk, 2019). One factor that may underlie participatory propaganda is the social utility that persons can derive from beliefs, even if they are false, which may stimulate them into rationalizing belief in falsehoods (Williams, 2022). The converse may also occur, where members of the public accurately report an experience, which is then taken up by others, usually political operatives or elites, and redeployed for a malign purpose. For example, technical problems with some voting machines in Arizona in 2022 were seized on by Trump and his allies as being an attempt to disenfranchise conservative voters (Reid, 2022). Both cases underscore the importance of audience involvement and the reverberating feedback loops between political actors and the public which can often amplify and extend the reach of intentional disinformation (Starbird et al., 2023; Vosoughi et al., 2018), and which

can often involve non-epistemic but nonetheless rational choices (Williams, 2021, 2022).

The circular and mutually reinforcing relationship between political actors and the public was a particularly pernicious aspect of the rhetoric associated with Trump’s big lie (for a detailed analysis, see Starbird et al., 2023). During the joint session of Congress to certify the election on 6 January 2021, politicians speaking in support of Donald Trump and his unsubstantiated claims about election irregularities appealed not to evidence or facts but to public opinion. For example, Senator Ted Cruz cited a poll result that 39% of the public believed the election had been “rigged”. Similarly, Representative Jim Jordan (R-Ohio), who is now Chairman of the House Judiciary Committee, argued against certification of the election by arguing that “80 million of our fellow citizens, Republicans and Democrats, have doubts about this election; and 60 million people, 60 million Americans think it was stolen” (Salek, 2023). The appeal to public opinion to buttress false claims is cynical in light of the fact that public opinion was the result of systematic disinformation in the first place. While nearly 75% of Republicans considered the election result legitimate on election day, this share dropped to around 40% within a few days (Arceneaux and Truex, 2022), coinciding with the period during which Trump ramped up his false claims about the election being stolen. By December 2020, 28% of American conservatives did not support a peaceful transfer of power (Weinschenk et al., 2021), perhaps the most important bedrock of democracy. Among liberals, by contrast, this attitude was far more marginal (3%).

Public opinion has shifted remarkably little since the election. In August 2023, nearly 70% of Republican voters continued to question the legitimacy of President Biden’s electoral win in 2020. More than half of those who questioned Biden’s win believed that there was solid evidence proving that the election was not legitimate (Agiesta and Edwards-Levy, 2023). However, the purported evidence marshaled in support of this view has been repeatedly shown to be false (Canon and Sherman, 2021; Eggers et al., 2021; Grofman and Cervas, 2023).¹ It is particularly striking that high levels of false election beliefs are found even under conditions known to reduce “expressive responding”—that is, responses that express support for a position but do not reflect true belief (Graham and Yair, 2023).

The entrenchment of the big lie erodes the core of American democracy and puts pressure on Republican politicians to cater to antidemocratic forces (Arceneaux and Truex, 2022; Jacobson, 2021, 2023). It has demonstrably decreased trust in the electoral system (Berlinski et al., 2021), and a violent constitutional crisis has been identified as a “tail risk” for the United States in 2024 (McLauchlin, 2023). Similar crises in which right-wing authoritarian movements are dismantling democratic institutions and safeguards have found traction in many countries around the world including liberal democracies (Cooley and Nexon, 2022).

In this context, it is worth noting that the situation in other countries, notably in the Global South, may differ from the situation in the U.S. (Badrinathan and Chauchard, 2024). On the one hand, low state capacity and infrastructure constraints may curtail the ability of powerful actors to spread disinformation and propaganda (though see Kellow and Steeves, 1998; Li, 2004, for discussion of the role of government-adjacent radio station RTLM in facilitating the 1994 Rwandan genocide). On the other hand, such spread can be facilitated by the fact that closed, encrypted social-media channels are particularly popular in the Global South, sometimes providing an alternative source of news when broadcast channels and other conventional media have limited reach. In those cases, dissemination strategies will also be less direct, relying more on distributed “cyber-armies” than direct one-to-millions broadcasts such as Trump’s

social-media posts (Badrinathan, 2021; Jalli and Idris, 2019). The harm that can be caused by such distributed systems was vividly illustrated by the false rumors about child kidnappers shared in Indian WhatsApp groups in 2018, which incited at least 16 mob lynchings, causing the deaths of 29 innocent people (Dixit and Mac, 2018). The ensuing interplay between the attempts of the Indian government to hold WhatsApp accountable and Meta, the platform's owner, highlights the limited power that governments in the Global South hold over multinational technology corporations (Arun, 2019). As a result, many platforms do not even have moderation tools for problematic content in popular non-Western languages (Shahid and Vashistha, 2023).

The power asymmetry between corporations and the Global South has been noted repeatedly, and recent calls for action include the idea of collective action by countries in the Global South to insist on regulation of platforms (Takhshid, 2021). We have only scratched the surface of a big global issue that is in urgent need of being addressed.

Despite these differences between the Global North and South, beliefs in political misinformation can be pervasive regardless of regime type or development level (e.g., for a discussion in the context of the “developing democracy” of Brazil, see Dourado and Salgado, 2021; Pereira et al., 2022).

The political landscape of disinformation

Given that the 2020 election was lost by the Republican candidate, the finding that conservatives are more likely than liberals to believe false election claims is explainable on the basis of motivated cognition and the general finding that conspiracy theories “are for losers” (Uscinski and Parent, 2014); that is, they provide an explanation—even if only a chimerical one—for a political setback to the losing parties. There is no a priori reason to assume that susceptibility to disinformation is skewed across the political spectrum.

However, a large body of recent research on the American public and U.S. political actors has consistently identified a pervasive ideological asymmetry, with conservatives and people from the populist right being far more likely to consume, share, and believe false information than their liberal counterparts (Benkler et al., 2018; Garrett and Bond, 2021; González-Bailón et al., 2023; Grinberg et al., 2019; Guess et al., 2020a; Guess et al., 2020b; Guess et al., 2019; Ognyanova et al., 2020). Research into the asymmetry culminated in a recent analysis of the news diet of 208 million Facebook users in the U.S., which discovered that a substantial segment of the news ecosystem is consumed exclusively by conservatives and that most misinformation exists within this ideological bubble (González-Bailón et al., 2023). Although the reasons for this asymmetry are not fully understood, Lasser et al. (2022) recently showed that it also held for politicians, with Republican members of Congress disseminating far more low-quality information on Twitter/X than their Democratic counterparts. Greene (2024) reported a parallel analysis for Facebook and found the same asymmetry between politicians of the two major parties. Similarly, Benkler et al. (2018) showed how the particular structure of the American media scene, with a dense interconnected cluster of right-wing sources that is separate from the remaining mainstream, fosters political asymmetry in the use and consumption of disinformation.

This asymmetry extends beyond the political domain to health-related information, which might at first glance appear to be of sufficient importance for most people to cast aside their political leanings. A recent systematic review discovered eight studies that identified conservatism as a predictor of susceptibility to health misinformation, seven studies that found no association involving political leanings, and not a single study that showed liberals to be more misinformed on health topics than conservatives (Nan

et al., 2022). The observed political asymmetry is also not limited to survey results or other behavioral measures. Wallace et al. (2023) examined vaccination and mortality data from two U.S. states (Ohio and Florida) during the COVID-19 pandemic and found a widening partisan gap in excess mortality. Specifically, whereas mortality rates were equal for registered Republican and Democratic voters pre-pandemic, a wide partisan gap—with excess death rates among Republicans being up to 43% greater than among Democratic voters—was observed after vaccines had become available for everyone. The gap was greatest in counties with the lowest share of vaccinated people and it almost disappeared for the most vaccinated counties. Similar results have been reported across U.S. states (Leonhardt, 2021). One explanation for these patterns invokes the frequent false statements by Republican politicians and conservative news networks—foremost Fox News—that discredited the COVID-19 vaccines (Hotez, 2023). In support, consumption of Fox News has been causally linked to lower vaccination rates (Pinna et al., 2022).

Moreover, a recent analysis identified a specific “Trump effect” such that even conditional on the Republican vote share, support for Trump was additionally and causally associated with a lower vaccination rate (Jung and Lee, 2023).

The political asymmetry surrounding the dissemination and consumption of misinformation must be caveated in two ways. First, although the asymmetry is substantial and pervasive it is not absolute. For some materials, such as specific conspiracy theories, the asymmetry is found to be attenuated in some studies (A. Enders et al., 2022; M. Enders and Uscinski, 2021). Second, the asymmetry observed among American politicians does not necessarily hold in other countries. Lasser et al. (2022) examined tweets by British and German parliamentarians and showed that with the exception of the extreme right in Germany (the AfD party), politicians across the mainstream spectrum were equally judicious in what information they shared in their tweets. This finding suggests that it is not conservatism per se that is associated with asymmetric reliance on misinformation, but the specific manifestation of conservatism currently dominant in the American political landscape.

Notwithstanding those caveats, the political asymmetry surrounding the dissemination and consumption of misinformation in the U.S. has been accompanied by at least two major issues: First, there has been a strong political response by Republicans in Congress who have commenced a campaign against misinformation research and researchers, claiming that the research seeks to censor conservative voices. Second, the political backlash has coincided with growing self-reflection and critique among scholars, some of whom began to question the misinformation research effort, culminating in claims that misinformation may not be sufficiently identifiable or widespread to warrant concern or countermeasures. We now take up these two issues in turn.

The politicization of misinformation research

At the time of this writing, Representative Jim Jordan, R-Ohio, has been leading a campaign against misinformation research and misinformation researchers in his role as Chairman of the House Judiciary Committee. The core allegation by Jordan and his allies² is that misinformation researchers are part of a purported “Censorship Industrial Complex” that is assisting the Biden administration in its purported endeavor to pressure platforms into suppressing conservative viewpoints (U.S. House of Representatives Judiciary Committee, 2023). The allegation is, however, problematic for at least four reasons: it rests on false assertions; it ironically denies first-amendment rights to researchers; it rests on a basic premise that is false; and it misunderstands the role of platforms in content moderation.

Concerning the first point, Jordan has subpoenaed several prominent academics engaged in the study of mis- and disinformation based on false assertions. For example, Dr Kate Starbird, an expert on disinformation from the University of Washington, was called to testify before Jordan's subcommittee and had to defend herself against accusations that she was colluding with the Biden administration in an effort to chill conservative speech (Nix and Menn, 2023). Core to the specific allegations against Starbird and her colleagues is a claim—initially voiced by online conspiracy theorists—that they colluded with the Department of Homeland Security to censor 22 million tweets during the 2020 election campaign. In actual fact, the researchers collected 22 million tweets for analysis, and flagged about 3000 (0.0001 of the total) for potential violations of Twitter's terms of use (Blitzer, 2023).

Second, Jordan's purported championing of free speech is difficult to reconcile with the chilling effect the House Committee's actions have had on the first-amendment rights of researchers. According to Starbird, "The people that benefit from the spread of disinformation have effectively silenced many of the people that would try to call them out" (Rutenberg and Myers, 2024). The deterring effect on the research community is widespread (Bernstein, 2023; Nix et al., 2023). Similarly, Facebook and YouTube have reversed their restrictions on content claiming that the 2020 election was stolen. Election disinformation, unsurprisingly, has seen an uptick in response (Rutenberg and Myers, 2024).

Third, Jordan's campaign rests on a false premise, namely that social-media platforms are biased against conservatives. Together with other conservative figures such as Tucker Carlson (formerly with Fox News) and Ben Shapiro, Jordan claimed in 2020 that "Big Tech is out to get conservatives". This claim has been shown to be wrong by several studies. For example, an analysis of Facebook engagements during the 2016 election campaign revealed that conservative outlets (Fox News, Breitbart, and Daily Caller) amassed 839 million interactions, dwarfing more centrist outlets (CNN with 191 million and ABC news with 138 million), and totaling more than the remaining seven mainstream pages in the top 10 (Barrett and Sims, 2021). Another analysis involving millions of Twitter users and 6.2 million news articles shared on the platform also found that conservatives enjoy greater algorithmic amplification than people on the political left (Huszár et al., 2022). Moreover, the Congressional January 6th Committee detailed the way in which major platforms, including Twitter and Facebook, facilitated the organization of the violent insurrection in a 122-page memo, although much of that information did not make it into the final committee report (Zakrzewski et al., 2023). Congressional investigators discovered that the platforms failed to heed their own experts' warnings about violent rhetoric on their platforms, and selectively failed to enforce existing rules to avoid antagonizing conservatives for fear of reprisals (Zakrzewski et al., 2023).

Finally, and perhaps most important, Jordan's pursuit fails to differentiate between the roles of government and the platforms, and in particular ignores the crucial role that platforms already play in shaping people's information diet (Lewandowsky et al., 2023a). In a nutshell, the internet is currently neither unregulated nor is all information on the internet equally free. Instead, nearly all content on social media is curated by algorithms that are designed to maximize dwell time in pursuit of the platforms' advertising profit (Lewandowsky and Pomerantsev, 2022; Wu, 2017). Algorithms therefore favor captivating information that keeps users engaged. Unfortunately, human attention is known to be biased towards negative information (Soroka et al., 2019), which creates an incentive for platforms to drench users in outrage-evoking content. Similar to junk food that supermarkets strategically place at checkout lanes, the information that is preferentially curated by platforms may satisfy

our presumed momentary preferences while reducing our long-term well-being. If platforms were to address their role in those dynamics, for example by redesigning their algorithms, this would hardly constitute censorship. Solving a problem one has caused is good iterative design rather than bias or suppression of opinions. No one would accuse a supermarket of suppressing consumers' preferences if the checkout lanes put on offer celery instead of chocolate bars.

In summary, far from being a restorative effort in defense of free speech, Jordan's attacks are reminiscent of similar campaigns launched against inconvenient scientists by the tobacco and fossil-fuel industries (Lewandowsky et al., 2023b). In all cases scientists have been subject to personal abuse, their email correspondence is hacked or subpoenaed, and allegations are woven together from snippets of decontextualized actions or events (Blitzer, 2023). Because these attacks are systemic, the response also requires a systemic approach (Desikan et al., 2023). However, any such response seems unlikely to be achievable in the current political landscape. Scientists who work under such challenging conditions must therefore rely on other avenues to protect their integrity. The U.S. National Academy of Sciences has published a list of resources for scientists under attack.³ Specific recommendations include responding publicly to valid criticism (without, however, engaging in a long drawn-out direct conversation with an attacker), reporting abusive messages to the authorities, and seeking support from colleagues who have been in similar situations (Kinser, 2020).

The attacks have also coincided with moves by the platforms and the courts that align with Jordan's claims. For example, the major platforms (Meta, Google, Twitter/X, and Amazon) have cut back on the number of staff dedicated to combating hate speech and misinformation (Field and Vanian, 2023). Meta (the parent company of Facebook) has been laying off employees in its "content review" team, which had been involved in countering misinformation and disinformation in the 2022 midterm election, citing confidence in improved electronic tools for detecting inauthentic accounts. It remains to be seen how the platform actions will play out during the 2024 presidential election.

In the legal arena, a Trump-appointed federal judge in Louisiana barred the Biden administration from having any contact with social-media companies and certain research institutions to discuss safeguarding elections in July 2023. The judgement echoed the claims by Jim Jordan and other Republicans that there was collusion between the White House and the social-media companies to censor conservative voices under the guise of fighting disinformation about COVID-19 during the pandemic and false election claims during the 2022 midterms. Although there are important and potentially problematic implications for free speech that arise whenever a government gets involved in managing what it considers misinformation (Neo, 2022; Vese, 2022), the Louisiana ruling was particularly broad in its prohibitions (West, 2023). The implications of the ruling include denying election officials access to information gathered by independent research bodies (the ruling lists "the Election Integrity Partnership, the Virality Project, the Stanford Internet Observatory, or any like project or group") that would enable them to debunk false election-related information and provide more accurate information instead. The Supreme Court blocked the Louisiana ruling in October 2023 (Hurley, 2023) but agreed to a full hearing later in its current term. We return to the conflict between free speech and the adverse effects of disinformation later.

The post-modern critiques of misinformation research

At the heart of research on misinformation is the belief that the concepts of truth and falsehood are essential to democracy, to cognition, and to daily life, and that the status of many, but of course not all claims, can be determined with sufficient accuracy

to warrant rebuttal of false information. For example, the “big lie” about a stolen election is just that—it is a lie with no sustainable evidentiary support and it is routinely referred to as such in the scholarly literature (e.g., Arceneaux and Truex, 2022; Canon and Sherman, 2021; Graham and Yair, 2023; Henricksen and Betz, 2023; Jacobson, 2021, 2023; Painter and Fernandes, 2022). The lie has been rejected by 62 American courts, all of which dismissed or ruled against law suits questioning the legitimacy of the election by Donald Trump or his supporters.⁴

It is curious that the reaction by Trump and some of his most ardent public supporters to such determinative judgments about the falsity of his claims has not been to claim that they are in fact true, but to attack the idea that objective knowledge is even possible. When confronted with a lie, Trump’s adviser Kellyanne Conway once famously quipped that she was presenting “alternative facts.” On another occasion, Trump’s attorney Rudy Giuliani declared that “truth isn’t truth.” Such a strategy seems oddly reminiscent of the postmodernist critique of the possibility of objective knowledge, which first arose as a core aspect of 1930s fascism and was then adapted by left-wing literary criticism from the 1960s onward (Lewandowsky, 2020). At that time, humanities scholars had grown increasingly uncomfortable with the idea that facts were just facts, and that there was no role for considering the personal or political interests of those who were engaged in the pursuit of empirical knowledge. In this, postmodernists raised an important point of self-reflection for scientists and others who blithely claimed that there was an impenetrable wall between facts and values. But then they took things too far. Derrida claimed that there was no such thing as objective knowledge. Foucault went on to suggest that—given this—all knowledge claims were nothing more than an assertion of the political interests of the investigator (McIntyre, 2018, p. 124).

This led to the “science wars” of the 1990s, when scientists and their allies fought back against subjectivism and relativism to defend the importance of objective knowledge at least as a regulative ideal of empirical inquiry. This particular attack on science eventually dissipated—and in the face of damage it had done to objective knowledge claims like the reality of global warming, some postmodernists such as Bruno Latour eventually even apologized (Latour, 2004)—but the damage was already done. Meanwhile, both the corporate sector and the religious and political right wing had once again taken up the strategy in their attacks on science. The advantage of post-modernism for anti-democratic purposes is obvious, and has echoes of authoritarian attacks on truth-tellers and their defenders throughout history. Indeed, to someone who embraces the idea that their political ideology *should* have supremacy over objective reality, the advantages of postmodernism are clear. Not only can falsehoods about the economy, crime, and political violence be offered as “alternative narratives” to carefully-measured statistics or other forms of evidence, but the credibility of *any* party as an objective truth-teller can be undermined. And this suits the authoritarian just fine—for where there is no truth then there can be no blame or accountability either.

Hannah Arendt long ago recognized the dangers of this strategy when she wrote: “the ideal subject of totalitarian rule is not the convinced Nazi or the convinced communist, but people for whom the distinction between fact and fiction ... true and false ... no longer exist.” This easy political slide into post-modernism does violence to the idea that truth matters, that facts can be discovered through empirical analysis, and that it is crucial to attempt to discern the facts before we can make good policy—especially when we hold competing values that will impact policy choice. And this is true even more so in an era when the creation and amplification of knowledge claims are so easily subject to digital manipulation and weaponization by anyone who has a

personal or political interest. Fortunately, researchers have developed conceptual, cognitive, and computational tools that permit the differentiation between legitimate contestation of facts on the one hand, and misinformation and willful disinformation on the other.

The identifiability of contested facts

Notwithstanding our rejection of the postmodernist project, we do not dispute its core idea that many contested assertions cannot be unambiguously adjudicated by referring to “facts”. There are indeed cases in which different actors may legitimately question each other’s “facts”. In our view, these ambiguous cases are precisely those that merit democratic debate and contestation. When conducted in good-faith, such debates can be particularly revealing because both sides can marshal evidence in support.

To illustrate, consider the recent controversy surrounding a machine-learning tool known as COMPAS (Dressel and Farid, 2018), which is intended to assist judges in the U.S. by predicting the likelihood of recidivism of a specific offender. Critics accused COMPAS of being racially biased based on statistical analysis of the evidence (Angwin et al., 2016). The case rested on the observation that among defendants who ultimately did not re-offend, the algorithm misclassified African-Americans as being at risk of re-offending more than twice as often as White offenders. This misclassification can have serious consequences for a person because judges are inclined to treat high-risk defendants more harshly.

Proponents of COMPAS rejected this charge and argued that the algorithm was not racially biased because it predicted recidivism equally for Black and White offenders for each of its 10 risk categories. That is, the classification into risk categories based on a large number of indicator variables was racially unbiased—a Black person’s actual probability of re-offending was the same as that of a White person with the same risk score (Dieterich et al., 2016).

It turns out that it is mathematically impossible to simultaneously satisfy both forms of fairness—calibration and classification—when the base rates of re-offending differ between groups (Berk et al., 2021; Lagioia et al., 2023). That is, if a greater share of Black people are classified as high-risk—which the algorithm does in an unbiased manner—then it necessarily follows that a greater share of Black defendants *who do not re-offend* will also be mistakenly classified as high-risk. In those circumstances, it would be inappropriate to accuse one or the other side of spreading misinformation, as each party has mathematical justification for their position and a resolution can only be attained through a value-laden policy discussion. Indeed, to our knowledge, the main contestants in this debate—Northpointe, the manufacturer of COMPAS (Dieterich et al., 2016) and ProPublica, a public-interest media organization (Angwin et al., 2016)—did not level charges of misinformation against each other despite engaging in robust debate.

A similar controversy with even greater stakes arose in the context of the COVID-19 vaccine rollout in the U.S. in 2021. Unlike most other countries, which vaccinated their populations according to age alone—with the elderly being given highest priority because of their much higher mortality rate from COVID-19—the U.S. Advisory Committee on Immunization Practices (ACIP) favored a policy that gave higher priority to essential workers (e.g., food and transport workers) than the elderly. This policy was partially motivated by the fact that racial minorities (Blacks and Hispanics) are underrepresented among adults over 65, whereas they are slightly over-represented among essential workers—thus, under an age-based policy the share of Whites who receive the vaccine would have initially been greater

than their proportion in the population would have warranted. Conversely, Blacks would have been underrepresented among the vaccinated early on (Mouk, 2023). This inequity could be avoided by first vaccinating essential workers among whom racial minorities were over-represented. However, because the age distribution of essential workers has a much lower average, fewer lives were saved among vaccinated essential workers—whose young age rendered their risk of dying from COVID-19 low to begin with—than would have been saved among the elderly had they been vaccinated (Rumpler et al., 2023). Modeling has confirmed that while the essential-worker policy introduced racial equity in terms of doses administered, more lives would have been saved *in all ethnic groups* under an age-based policy (Rumpler et al., 2023). Again, the apparent fairness of a policy depended on the outcome measure: doses administered vs. lives saved. Given the unequal distributions of different ethnic and racial groups across different ages, no mathematical possibility exists to settle on a single “fair” policy. Public opinion appears to have been broadly in line with the policy ultimately adopted by ACIP (Persad et al., 2021).

The controversies surrounding COMPAS and ACIP’s vaccination policy are just two instances of a much wider problem, which is that when issues become sufficiently complex, even good-faith actors may find it impossible to agree. One reason is that cognitive limitations prevent a full Bayesian representation (the gold standard of rationality) of the problem (Pothos et al., 2021). Instead, people are forced to simplify their representations, for example by partitioning their knowledge (Lewandowsky et al., 2002). Persistent and irresolvable disagreements are thus almost ensured by human cognitive limitations (Pothos et al., 2021). The second reason is that people differ in their values and weigh evidence differently even if all parties can agree on underlying facts (Walasek and Brown, 2023).

Nonetheless, controversies such as those surrounding COMPAS and ACIP’s vaccination policy do not give licence to political actors to obscure the debate through falsehoods, misleading claims, or lies. On the contrary, proper debate of those issues is only possible in the absence of falsehoods because their resolution ultimately requires a trade-off of values that is best arrived at by weighing the importance of different competing sets of evidence. We therefore reject recent academic voices that have questioned whether misinformation can be reliably identified at all (Acerbi et al., 2022; Adams et al., 2023; Harris, 2022; van Doorn, 2023; Yee, 2023a, 2023b). We suggest that its identification is essential and, as we show next, empirically well supported.

The identifiability of misinformation

We place our case into the context of the more extreme end of the academic critique because it involves positions that are antithetical to ours, calling into question the entire idea of fact-checking. For example, Uscinski (2015) raised the specter that fact-checking is merely a “veiled continuation of politics by means of journalism” (p. 243). Yee (2023a) argued more broadly that any deference to “epistemic elites”—including not only fact-checkers but also academics, researchers, or journalists—is problematic, and assessment of the quality of information should include democratic elements “that are participatory, transparent, and fully negotiable by average citizens” (Yee, 2023a, p. 1111). This demand has several problematic implications. First, it does not explain who counts as “average citizen” and who would belong to the “elite”. At what point should individuals seeking to counter misinformation begin to recuse themselves for fear of accidentally treading on “average” citizens? Is a virologist too “elite” to correct misinformation surrounding the origin of a new virus? What about citizens with a PhD or Master’s degree, how

are they being classified? Second, why exactly would one exclude epistemic elites, such as investigative journalists or forensic IT experts, from identifying bad-faith actors such as foreign “bots” or “trolls”? Are average citizens really better at this task than network scientists? Should we decide by social-media poll whether a new strain of avian flu is contagious to humans (Lewandowsky et al., 2017)? Probably not. There are obviously many domains that benefit from expert assessment of claims.

Nonetheless, there has been much research that has revealed the competence of crowds in the context of fact-checking. For example, Pennycook and Rand (2019) showed that crowdsourced trust ratings of media outlets were quite successful in the aggregate when compared to ratings by professionals, notwithstanding substantial partisan differences. This basic finding has been replicated and extended several times (M. R. Allen et al., 2024; Martel et al., 2024), with community-based fact-checking of COVID-19 content being 97% accurate in one study (M. R. Allen et al., 2024). Care must, however, be taken that crowds are politically balanced. When people can choose what content to evaluate, as in Twitter/X’s crowdsourced “Birdwatch” fact-checking program (now known as Community Notes), partisan differences among contributors may limit the value of the crowdsourcing (J. Allen et al., 2022). The crowdsourcing results show not only that average citizens can match the competence of experts in the aggregate, but they also reaffirm that misinformation is identifiable.

Much recent research has uncovered specific “fingerprints” that can enable people as well as machines to infer the likely quality or accuracy of content. Misinformation has been shown to be suffused with emotions, logical fallacies, and conspiratorial reasoning (Blassnig et al., 2019; Carrasco-Farré, 2022; Fong et al., 2021; Musi et al., 2022; Musi and Reed, 2022). For example, critical thinking methods offer a qualitative approach to deconstructing arguments in order to identify the presence of reasoning fallacies (Cook et al., 2018).

Quantitatively, one study found that compared to reliable information, misinformation is less cognitively complex and 10 times more likely to rely on negative emotional appeals (Carrasco-Farré, 2022). In confirmation, numerous other studies show that misinformation is, on average, more emotional than factual information (for a systematic review, see Peng et al., 2023) Upward of 75% of anti-vaccination websites use negative emotional appeals (Bean, 2011) and linguistic analyses show that conspiracy theorists use significantly more fear-driven language as compared to scientists (Fong et al., 2021).

Emotion also plays a role in the receivers’ behavior. People have been shown to be more susceptible to misinformation when put in an emotional state (Martel et al., 2020), which helps explain the preferential and more rapid diffusion of unreliable versus reliable information online (Pröllochs et al., 2021; Vosoughi et al., 2018).

Critics may argue that the datasets used for determining what constitutes “misinformation” and “reliable” information are limited or biased or that the mere prevalence of these cues is not evidence of their diagnosticity in real-world contexts. However, computational machine-learning work relying on a large variety of different URL sources and fact-checked datasets has confirmed that the results are robust and generalizable (Ghanem et al., 2020; Kumari et al., 2022; Lebernegg et al., 2024). A recent comprehensive study which combined many of the available cues found that they have high diagnostic and predictive validity and help discriminate between false and true information, with state-of-the-art models reaching over 83% classification accuracy (Lebernegg et al., 2024). Moreover, real-world training on fake news detection, such as logical fallacy training, helps people accurately discriminate between misleading and credible news

(e.g., Hruschka and Appel, 2023; Lu et al., 2023; Roozenbeek et al., 2022).

In summary, the available evidence shows quite convincingly that misinformation can be identified by both humans and machines with considerable accuracy. As we show next, we can go beyond mere identification as there are also at least three ways in which one can ascertain the deceptive intent underlying disinformation if present. Identification of deceptive intent is particularly pertinent because it allows information to be safely discounted without requiring a detailed analysis of its factual status.

The identifiability of willful disinformation

For decades, the hallmark of Western news coverage about politicians' false or misleading claims was an array of circumlocutions that carefully avoided the charge of lying—that is, knowingly telling an untruth with intent to deceive (Lackey, 2013)—and instead used adverbs such as “falsely”, “wrongly”, “bogus”, or “baseless” when describing a politician’s speech. Other choice phrases referred to “unverified claims” or “repeatedly debunked claims”. This changed in late 2016, when the *New York Times* first used the word “lie” to characterize an utterance by Donald Trump (Borchers, 2016). The paper again referred to Donald Trump’s lies within days of the inauguration in January 2017 (Barry, 2017) and it has grown into a routine part of its coverage from then on. Many other mainstream news organizations soon followed suit and it has now become widely accepted practice to refer to Trump’s lies as lies.

Given that lying involves the intentional uttering of false statements, what tools are at our disposal to infer a person’s intention when they utter falsehoods? How can we know a person is lying rather than being confused? How can we infer intentionality?

Anecdotally, defenders of Donald Trump’s lies have raised precisely that objection to the use of the word “lie” in connection with his falsehoods. This objection runs afoul of centuries of legal scholarship and Western jurisprudence. Brown (2022) argues that inferring intentionality from the evidence is “ordinary and ubiquitous and pervades every area of the law” (p. 2). Inferring intentionality is the difference between manslaughter and murder and is at the heart of the concept of perjury—namely, willfully or knowingly making a false material declaration (Douglis, 2018).

There are at least three approaches that can be pursued to infer intentional deception by a communicating agent with varying degrees of confidence. The first approach is statistical and relies on linguistic analysis of material. Unlike people, who are not very good lie detectors despite performing (slightly) above chance (Bond and DePaulo, 2006; Mattes et al., 2023), recent advances in natural language processing (NLP) have given rise to machine-learning models that can classify texts as deceptive or honest based on subtle linguistic clues (e.g., Braun et al., 2015; Davis and Sinnreich, 2020; Van Der Zee et al., 2021). To illustrate, a model that relied on analysis of the distribution of different types of words achieved 67% accuracy (considerably better than the 52% achieved by human judges) on texts generated by speakers who were either instructed to lie or to be honest. Using the same analysis approach, Davis and Sinnreich (2020) trained a model to classify tweets by Donald Trump as true or false by using independent fact-checks as ground truth. The model was able to classify tweets with more than 90% accuracy, suggesting that Trump uses subtly different language (e.g., more negative emotion, more prepositions and discrepancies) when communicating untruths. A similar model of Trump’s tweets was developed by Van Der Zee et al. (2021), who additionally applied 26 extant models from the literature to Trump’s tweets and showed that most of them performed above chance despite being developed on

very different materials. In summary, NLP-based approaches have repeatedly shown their value in the classification of speech into honest and deceptive. The fact that those models succeed also when applied to the tweets of Donald Trump implies at the very least that Trump’s falsehoods are not uttered at random or accidentally but are deployed using specific linguistic techniques.

In general, machine-learning approaches to deception detection have shown promise. A recent systematic review identified 81 studies, 19 of which achieved accuracies in excess of 90%, with a further 15 exceeding 80% accuracy (Constâncio et al., 2023). The machine-learning models in that ensemble were trained on a variety of corpora, ranging from reviews on Tripadvisor (either true or generated with the intent to deceive; Barsever et al., 2020) to segments of a radio game show dedicated to bluff detection by the audience (Papantoniou et al., 2021). In all cases, the ground truth (i.e., whether or not deceptive intent was present) was unambiguously known, and the models learned to identify deceptive text based on linguistic analysis with considerable albeit imperfect success.

The second approach to establish willful deception relies on analysis of internal documents of institutions such as governments or corporations. Comparison of the internal knowledge to public stances of the same entities can identify active deception, especially when it is large-scale. Numerous such cases exist, mainly involving corporations and their associated infrastructure such as think-tanks and other front groups (Ceccarelli, 2011; Oreskes and Conway, 2010). For example, as early as the 1920s, the electricity industry organized a propaganda campaign to falsely insist that private sector electricity was cheaper and more reliable than electricity generated in the public sector (Oreskes and Conway, 2023). The tobacco industry’s activities to mislead the public about the dangers from smoking are well documented and established beyond reasonable doubt (e.g., Cataldo et al., 2010; Fallin et al., 2013; Francey and Chapman, 2000; Proctor, 2012). The tobacco industry was well aware of the link between smoking and lung cancer in the 1950s and 1960s (Proctor, 2012), and yet continued publicly to dispute that medical fact using a variety of propagandistic means (Landman and Glantz, 2009; Proctor, 2011). Similarly, analysis of internal documents of the fossil-fuel industry has revealed that industry leaders, in particular ExxonMobil, were fully aware of the reality of climate change and its underlying causes (Supran and Oreskes, 2017, 2021) while simultaneously expending large sums to deny its existence in public (J. Farrell, 2016) and to prevent Congress from enacting climate-mitigation legislation (Brulle, 2018). Ironically, ExxonMobil’s scientists projected global temperatures in the 1970s and 1980s with comparable skill as independent academics at the time (Supran et al., 2023). As Baker and Oreskes (2017) argued, the best explanation for ExxonMobil’s conduct is that they knowingly deceived the public by funding a disinformation machine that denied the realities of climate change. This approach admittedly requires considerable resources and skill, and it is comparatively slow, but in exchange the results it yields are particularly diagnostic and demonstrably useful in litigation. In the case of the tobacco industry, this was the basis for a conviction of Phillip Morris under federal racketeering (RICO) law. The appeals in that case explicitly noted that Phillip Morris intentionally deceived the public and that first-amendment (free speech) rights did not apply as they do not protect fraud or deliberate misrepresentation (Farber et al., 2018). In the case of the fossil fuel industry, litigation has not been met with notable success at the time of this writing, but the “Exxon knew” campaign, based on research by Supran and colleagues (Supran et al., 2023; Supran and Oreskes, 2017, 2021), has had considerable public impact with 178 relevant media articles identified by Google News.⁵

The final approach to identifying intentional deception resembles the approach involving institutional documents but specifically focuses on lies promulgated by identifiable individuals. We illustrate this approach with Donald Trump's big lie about the 2020 presidential elections, focusing on statements made in courts of law. Although Trump was making widespread public accusations of fraud, his lawyers—who filed more than 60 lawsuits in connection with the election—did not echo those accusations in court. Quite on the contrary, his lawyers frequently disavowed any mention of fraud in court despite their very different public stance. For example, Rudy Giuliani, one of Trump's lead attorneys, stood outside a landscaping business on the day most networks declared the election for Biden, and thundered that "It's [the election] a fraud, an absolute fraud." Ten days later, being questioned by a federal judge in Pennsylvania during one of Trump's lawsuits (dealing with whether local election officials in Pennsylvania should have allowed voters to fix problems with their mail-in ballots after submitting them), he declared "This is not a fraud case" (Lerer, 2020). This pattern was pervasive: Trump's lawyers continued to back away from suggestions that the election was stolen and admitted in court that there was no evidence of fraud, all in contradiction to their plaintiff's public statements (Lerer, 2020).

Notwithstanding the careful hedging of their claims in court, the frivolous suits filed on behalf of Trump resulted in sanctions for several of his attorneys. Two lawyers who did claim widespread voter fraud not only had their suit dismissed but were also sanctioned \$187,000 by a federal judge in Colorado for their frivolous, meritless case (Polantz, 2021). The decision was upheld on appeal, and the Supreme Court declined to hear a further appeal by the lawyers (Scarcella, 2023). Altogether, 22 Trump lawyers have been identified who face sanctions in litigation, criminal prosecutions, and state bar disciplinary proceedings. In all cases, what appears to be at issue is violation of the Model Code of Conduct, in particular rules stipulating that claims must be meritorious and that lawyers must exhibit candor and truthfulness (Neff and Fredrickson, 2023).

Since the flurry of lawsuits in late 2020, Trump lawyer Sidney Powell has pleaded guilty to charges arising from her involvement in pushing the big lie. Ms Powell pleaded guilty to "conspiracy to commit intentional interference with performance of election duties" and agreed to cooperate with prosecutors in a criminal case against Donald Trump (Fausset and Hakim, 2023). Two further Trump lawyers have pleaded guilty in the same case and agreed to testify truthfully about other defendants (Blake, 2023).

In a civil suit brought against Rudy Giuliani by two election workers in Georgia, whom he had publicly accused of election fraud, Giuliani conceded before trial that those statements were false (Brumback, 2023). The election workers were awarded \$148 million in damages, causing Giuliani to file for bankruptcy in late 2023 (Aratani and Oladipo, 2023). In a further twist, Giuliani repeated his false claims during the trial *outside the court room* even while his lawyers conceded in court that they were wrong (Hsu and Weiner, 2023).

Giuliani was promptly sued again by the election workers, and at the time of this writing the suit was still under way (Hsu and Weiner, 2023).

The big lie was not just curated and pushed by politicians seeking to cling to power and their attorneys. It is now public knowledge that one major news network, Rupert Murdoch's Fox News, knowingly amplified claims about the election that network executives knew to be false. The fact that Fox lied became apparent during a defamation suit filed by Dominion Voting Machines against the network over false allegations that the voting machines had been rigged to steal the 2020 election. As trial was about to begin, Fox News agreed to pay Dominion \$787.5 million and acknowledged that the network had broadcast false statements. The discovery process that preceded trial had

uncovered numerous documents and emails that revealed that senior network executives and hosts were convinced that the allegations about the election made by Trump and his allies were untrue (e.g., Peltz, 2023; Terkel et al., 2023). The network continued to air those allegations and its CEO instructed staff that fact-checking "had to stop" because it was bad for business (Levine, 2023). One scholar put it succinctly: "Fox News deliberately misleads the audience for profit" (Nyberg, 2023, p. 1). Although Fox has been repeatedly implicated in spreading disinformation with harmful consequences for the American public (Ash et al., 2023; Bolin and Hamilton, 2018; Bursztyrn et al., 2020; DellaVigna and Kaplan, 2007; Feldman et al., 2012; Kull et al., 2003; Simonov et al., 2022), the Dominion case provided a unique opportunity to ascertain that, at least in this case, the network was knowingly lying to its audience.

The preceding examples illustrate the approaches available to establish—with some degree of confidence—the intention to deceive that is the core element of lies. Our examples are not intended to be exhaustive but they illustrate the options available to researchers, journalists, and the public to uncover when they are being lied to. The examples also put to rest several generous auxiliary assumptions that have been made about lies in politics, such as their presumed inevitability because issues can be so nuanced that complete honesty is impossible. Contrary to that assumption, the fact that a person's rhetoric can differ strikingly between courts of law—where penalties apply for misrepresentations and perjury—and politics—where accountability is notoriously absent—not only reveals the intention to deceive but also the person's sensitivity to the consequences of their speech.

We have already noted that the contrast between what companies such as ExxonMobil or Philip Morris said in public about their products and what they discussed in private was sufficient to provoke legal consequences. Similar arguments, that fraudulent political speech should not be protected by the First Amendment, have been advanced in the context of Trump's big lie (Henricksen and Betz, 2023).

Although our examination was necessarily limited to a small number of cases, they suffice to illustrate a pathway towards pinpointing intentional disinformation by analysing the utterances of the liars themselves, be they corporations, politicians, or media organizations. We believe that the basic approach is of considerable generality, extending to numerous recorded instances:

- Politicians catching themselves lying by changing their story, indicating they were telling an untruth on at least one of those occasions (O'Toole, 2022, p. 427).
- Attorneys of conspiracy theorist Alex Jones—who was sued for his claims that the Sandy Hook massacre never happened by parents of the victims—seeking to defend him by calling him a performance artist who should not be taken seriously (Borchers, 2017).
- Alex Jones himself admitting in court that the Sandy Hook shooting was "100% real" after having misled millions of people for many years (Associated Press, 2022).
- Fox News requiring their employees to be vaccinated against COVID-19 or submit to daily testing while the network routinely broadcast anti-vaccination content (Darcy, 2021).
- Tucker Carlson, former Fox News host, openly admitting that he lies on air (Muzaffar, 2021).

Moving forward

Our work explored three fundamental premises: First, that democracy rests on a foundation of common knowledge (H. Farrell and Schneier, 2018) and that it is imperiled if citizens

cannot agree on basic facts such as the integrity of elections (H. Farrell and Schneier, 2018; Tenove, 2020). Second, that while democratic debate—including evidence-informed policy-making—often involves contestation of facts (e.g., Kuklinski et al., 1998), this does not licence the use of outright lies and propaganda to willfully mislead the public (Lewandowsky, 2020). Third, that it is often possible to identify falsehoods, disinformation, and lies and differentiate them from good-faith political and policy-related argumentation.

At the time of this writing, Donald Trump is the Republican nominee for the 2024 presidential election. His campaign has rolled out an explicitly authoritarian agenda for his second term (Arnsdorf and Stein, 2023). The authoritarian agenda is likely to result in less free speech, rather than more, which is ironic in light of the fact that people such as Jim Jordan, who are attacking the idea of studying disinformation, do so under the banner of defending the First Amendment. Against this background, the question of how to address Donald Trump's lies in particular and misinformation in general takes on particular importance.

At the more pessimistic end, Barkho (2023) posed three questions about the success of fact-checking Trump's claims: first, have fact-checkers succeeded in persuading Trump to stop disseminating lies? Second, have the long inventories of falsehoods compiled by fact-checkers embarrassed or shamed Trump? Third, has fact-checking changed public perception of what constitutes truth? At first glance, the answer to all three questions might appear to be a resounding “no” (even though the counterfactual is, of course, unknown). However, at the more optimistic end of the spectrum, experimental studies in which election-fraud misinformation was corrected have found positive effects on trust in electoral processes (Bailard et al., 2022; Painter and Fernandes, 2022), including among Republican respondents and supporters of Trump. Those findings should give rise to a sliver of optimism that even partisans are receptive to corrective messages about election integrity, and therefore underscore the value of disinformation research.

Correcting lies about elections is arguably compatible with the spirit of a democracy. But what is the democratic legitimacy of broader countermeasures against misinformation and disinformation? It is straightforward to explore techniques with which to correct misconceptions in an experiment, in particular if the misinformation is introduced in the experiment itself (e.g., Ecker et al., 2011). It is less straightforward to deploy such techniques in the public sphere. Who determines what is “misinformation”, and what is “correct”? And how narrow is the gap between correcting misinformation and banning it? Several countries have recently outlawed “fake news” (e.g., Burkina Faso, Cambodia, Hungary, India, Malaysia, Singapore, and Vietnam) whose democratic credentials are at best questionable. In those cases, fake news can damage democracy not only by disinforming the public but also because countermeasures can be used to curb civil liberties and justify authoritarian crackdowns (Neo, 2022; Vese, 2022). Indeed, given that Donald Trump has routinely labeled any media coverage he did not like as “fake news”, perhaps the worst response to misinformation would be a law against fake news designed by Donald Trump and his allies.

There are, however, numerous ways in which the public can be better protected by the platforms—in particular if prodded into action by suitable regulations—against disinformation. One avenue involves content moderation and removal of unacceptable or problematic content, such as hate speech. The public is broadly supportive of moderation in certain cases (Kozyreva, Herzog, et al., 2023), and the European Union's recent Digital Services Act (DSA) acknowledges a role for content moderation while highlighting the need for transparency of the underlying rules (for details, see Kozyreva, Smillie, et al., 2023). In addition, there are a

number of alternative approaches that aim to inform or educate consumers rather than govern content directly. Those approaches have the advantage that they side-step concerns about censorship and that they are demonstrably scalable and readily deployable by the platforms.

One avenue involves the provision of “nutrition labels”, that is, indicators of the quality of a source. Reliable indicators of quality exist that are based on basic journalistic principles (Lin et al., 2023), and it is well-known that perceived source credibility can influence misinformation persuasiveness (Nadarevic et al., 2020; Prike et al., 2024). The effectiveness of source-quality indicators can be enhanced by introducing friction, for example, by requiring users to expend additional clicks to make information visible (L. Fazio, 2020; Pillai and Fazio, 2023). Naturally, such indicators cannot be perfect, and even sources of widely-acknowledged high quality can also publish dubious content. This makes it important to go beyond credibility and consider alternative approaches, such as those that boost users' ability to spot deception and enhance their information-discernment skills. This can range from teaching “critical ignoring” (Kozyreva, Wineburg, et al., 2023), which enables people to ignore information that is unlikely to warrant expenditure of our limited attention, to psychological inoculation or “prebunking” (Lewandowsky and van der Linden, 2021; Roozenbeek et al., 2022), which involves refuting a lie in advance by explaining the rhetorical techniques that disinformers use to mislead consumers (e.g., scapegoating, false dichotomies, ad hominem attacks, and so on). Through short “edutainment” videos that are displayed as ads or public-service messages, this approach has been scaled on social media to empower millions of people to spot manipulation techniques (Goldberg, 2023). Meta-analyses have affirmed the efficacy of the inoculation approach (Banas and Rains, 2010; Lu et al., 2023). However, while standard debunking and prebunking interventions promise to be effective regardless of the cultural context in which they are applied (Blair et al., 2024; Pereira et al., 2023; Porter and Wood, 2021; but see Pereira et al., 2022), the effects of other interventions such as media-literacy training may be less robust in the Global South (Badrinathan, 2021). Some interventions developed and successfully applied in the Global North may also be less suitable in less-developed countries, if for example they target dissemination channels that have limited relevance locally (Badrinathan and Chauchard, 2024; de Freitas Melo et al., 2019).

Overall, much is now known about various cognitively-inspired countermeasures to correct misinformation or to protect people against being misled in the first place. For further extensive discussion of these countermeasures, see Ecker et al. (2022) and Kozyreva et al. (2024). Some of the cognitive science of misinformation has been reflected in European regulatory initiatives, such as the strengthened Code of Practice on Disinformation (Kozyreva, Smillie, et al., 2023). In addition, specific evidence-based recommendations for platforms have been developed by Roozenbeek et al. (2023) and Wardle and Derakhshan (2017).

Our work has also identified several important questions for future research. We consider the long-term consequences of misinformation on society to be a particularly pressing issue. We have a reasonably good understanding of the individual-level cognitive processes that are engaged when a person is exposed to a single piece of misinformation (Ecker et al., 2022). We know very little about the cognitive and social consequences for an individual who is inundated with information of dubious quality for prolonged periods of time. We do not know how societies are affected by epistemic uncertainty and chaos in the long run. Numerous indicators suggest that Western societies, in particular the United States, are

ailing (e.g., Lewandowsky et al., 2017), but the attribution of those trends to misinformation or epistemic chaos is difficult. On those occasions where researchers have successfully isolated causal effects, they tend to implicate certain media organs (e.g., Fox News in particular) in compromising public health (Bursztyn et al., 2020; Simonov et al., 2020), and they have identified the role of social media in causing ethnic hate crimes and xenophobia (Bursztyn et al., 2019; Müller and Schwarz, 2021). However, it is unclear as yet how generalizable those findings are and much additional work remains to be done (for a review, see Lorenz-Spreen et al., 2022).

Future research should also address some of the limitations of fact-checking, such as the difficulties of verifying statements about the future (Nieminen and Sankari, 2021) or arguments that employ the rhetorical technique of “paltering” — that is, the use of truthful statements to convey a misleading impression (Lewandowsky et al., 2016; Rogers et al., 2017). One approach is to focus on what is pragmatically useful for people to make informed decisions, such as whether a claim is misleading (Birks, 2019), with critical thinking methods offering a means of identifying the presence of logical fallacies (Cook et al., 2018).

Increasing research attention is being paid to the concept of discernment; that is, the extent to which accurate misinformation is believed more than misinformation (Pennycook and Rand, 2021). Focusing on discernment rather than acceptance of misinformation guards against inadvertently developing interventions that reduce belief in facts and misinformation equally. A general cynicism and disbelief of everything does not solve the misinformation problem. Instead, we must boost people’s ability to distinguish between facts and falsehoods.

Conclusion

We began the paper with a quote from Hannah Arendt, one of the foremost analysts of 20th century totalitarianism. It is worth here revisiting the same quotation in its extended form, which underscores the urgency of finding a solution to the epistemic crisis affecting democracy in the U.S. and beyond:

“If everybody always lies to you, the consequence is not that you believe the lies, but rather that nobody believes anything any longer.... And a people that no longer can believe anything cannot make up its mind. It is deprived not only of its capacity to act but also of its capacity to think and to judge. *And with such a people you can then do what you please.*” (our emphasis)

— Hannah Arendt

Received: 25 January 2024; Accepted: 22 July 2024;

Published online: 31 July 2024

Notes

- Further detailed debunkings of election disinformation are provided by the Cybersecurity and Infrastructure Security Agency at <https://www.cisa.gov/topics/election-security/rumor-vs-reality>.
- We focus here on the activities of Jim Jordan because he is the acknowledged leader of a political counter movement aimed at misinformation research. This must not be taken to imply that Jordan is the only political actor involved in this effort.
- <https://www.nationalacademies.org/documents/embed/link/LF2255DA3DD1C41C0A42D3BEF0989ACAEC3053A6A9B/file/DC4CDD2AC5D4B2DB08255A7EA6244AA9D7CA6F951C22?noSaveAs=1>
- One ruling that was initially in Trump’s favor was later overturned by the Pennsylvania Supreme Court. Canon and Sherman (2021) provide a list of cases.
- Search conducted on 10 April 2024.

References

- Acerbi A, Altay S, Mercier H (2022) Research note: Fighting misinformation or fighting for information? *Harv Kennedy School Misinform Rev* 3. <https://doi.org/10.37016/mr-2020-87>
- Adams Z, Osman M, Bechliyanidis C, Meder B (2023) (Why) Is Misinformation a Problem? *Perspect Psychol Sci* 17456916221141344. <https://doi.org/10.1177/17456916221141344>
- Agiesta J, Edwards-Levy A (2023) CNN poll: Percentage of Republicans who think Biden’s 2020 win was illegitimate ticks back up near 70%. *CNN*. <https://edition.cnn.com/2023/08/03/politics/cnn-poll-republicans-think-2020-election-illegitimate/index.html>
- Allen J, Martel C, Rand DG (2022) Birds of a feather don’t fact-check each other: Partisanship and the evaluation of news in Twitter’s Birdwatch crowdsourced fact-checking program. *CHI Conf Hum Factors Comp Syst* 1–19. <https://doi.org/10.1145/3491102.3502040>
- Allen MR, Desai N, Namazi A, Leas E, Dredze M, Smith DM, Ayers JW (2024) Characteristics of X (formerly Twitter) Community Notes addressing COVID-19 vaccine misinformation. *JAMA* 331:1670. <https://doi.org/10.1001/jama.2024.4800>
- Angwin J, Larson J, Mattu S, Kirchner L (2016) Machine bias: There’s software used across the country to predict future criminals and it’s biased against blacks. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Aratani L, Oladipo G (2023) Giuliani files for bankruptcy after judge rules Georgia election workers can collect \$148m. *The Guardian*. <https://www.theguardian.com/us-news/2023/dec/21/giuliani-148-million-damages-georgia-lawsuit>
- Arceneaux K, Truex R (2022) Donald Trump and the Lie. *Perspect Polit* 1–17. <https://doi.org/10.1017/S1537592722000901>
- Arnsdorf I, Stein J (2023) Trump touts authoritarian vision for second term: ‘I am your justice’. *Washington Post*. <https://www.washingtonpost.com/elections/2023/04/21/trump-agenda-policies-2024/>
- Arun C (2019) On WhatsApp, rumours, and lynchings. *Econ Polit Wkly* 54(6):30–35
- Ash E, Galletta S, Hangartner D, Margalit Y, Pinna M (2023) The effect of Fox News on health behavior during COVID-19. *Polit Anal* 1–10. <https://doi.org/10.1017/pan.2023.21>
- Associated Press (2022) Alex Jones concedes that the Sandy Hook attack was ‘100% real’. *NPR*. <https://www.npr.org/2022/08/03/1115414563/alex-jones-sandy-hook-case>
- Badrinathan S (2021) Educative interventions to combat misinformation: evidence from a field experiment in India. *Am Polit Sci Rev* 1–17. <https://doi.org/10.1017/S0003055421000459>
- Badrinathan S, Chauchard S (2024) Researching and countering misinformation in the Global South. *Curr Opin Psychol* 55:101733. <https://doi.org/10.1016/j.copsyc.2023.101733>
- Bailard, CS, Porter, E, & Gross, K (2022). Fact-checking Trump’s election lies can improve confidence in U.S. elections: Experimental evidence. *Harv Kennedy School Misinform Rev*. <https://doi.org/10.37016/mr-2020-109>
- Baker E, Oreskes N (2017) Science as a game, marketplace or both: a reply to Steve Fuller. *Soc Epistemol Rev Reply Collect* 6:65–69
- Banas JA, Rains SA (2010) A meta-analysis of research on inoculation theory. *Commun Monogr* 77:281–311
- Barkho L (2023) A critical inquiry into US media’s fact-checking and compendiums of Donald Trump’s falsehoods and “lies”. In A Akande (Ed.) *The perils of populism: The end of the American century?* (pp. 259–278). Springer
- Barrett PM, Sims JG (2021) False accusation: The unfounded claim that social media companies censor conservatives (tech. rep.). *New York University Stern Center for Business and Human Rights*
- Barry D (2017). In a swirl of ‘untruths’ and ‘falsehoods,’ calling a lie a lie. *New York Times*. <https://www.nytimes.com/2017/01/25/business/media/donald-trump-lie-media.html>
- Barsever D, Singh S, Neftci E (2020) Building a better lie detector with BERT: The difference between truth and lies. 2020 International Joint Conference on Neural Networks (IJCNN). <https://doi.org/10.1109/ijcnn48605.2020.9206937>
- Bean SJ (2011) Emerging and continuing trends in vaccine opposition website content. *Vaccine* 29:1874–1880. <https://doi.org/10.1016/j.vaccine.2011.01.003>
- Benkler Y, Faris R, Roberts H (2018) *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press
- Berk R, Heidari H, Jabbari S, Kearns M, Roth A (2021) Fairness in criminal justice risk assessments: The state of the art. *Sociol Methods Res* 50:3–44. <https://doi.org/10.1177/0049124118782533>
- Berlinski N, Doyle M, Guess AM, Levy G, Lyons B, Montgomery JM, Nyhan B, Reifler J (2021) The effects of unsubstantiated claims of voter fraud on confidence in elections. *J Exp Polit Sci* 10(1), 34–49. <https://doi.org/10.1017/xps.2021.18>
- Bernstein A (2023) Republican Rep. Jim Jordan issues sweeping information requests to universities researching disinformation. *Pro Publica*. <https://www.propublica.org/article/jim-jordan-disinformation-subpoena-universities>

- Birks J (2019) Fact-checking journalism and political argumentation: A British perspective. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-30573-4>
- Blair RA, Gottlieb J, Nyhan B, Paler L, Argote P, Stainfield CJ (2024) Interventions to counter misinformation: Lessons from the Global North and applications to the Global South. *Curr Opin Psychol* 55:101732. <https://doi.org/10.1016/j.copsyc.2023.101732>
- Blake A (2023) Jenna Ellis's tearful guilty plea should worry Rudy Giuliani. *Washington Post*. <https://www.washingtonpost.com/politics/2023/10/24/jenna-ellis-guilty-plea-georgia-giuliani-trump/>
- Blassnig S, Büchel F, Ernst N, Engesser S (2019) Populism and informal fallacies: an analysis of right-wing populist rhetoric in election campaigns. *Argumentation* 33:107–136. <https://doi.org/10.1007/s10503-018-9461-2>
- Blitzer J (2023) Jim Jordan's conspiratorial quest for power. *The New Yorker*. <https://www.newyorker.com/magazine/2023/10/30/jim-jordans-conspiratorial-quest-for-power>
- Bolin JL, Hamilton LC (2018) The news you choose: News media preferences amplify views on climate change. *Environ Polit*. <https://doi.org/10.1080/09644016.2018.1423909>
- Bond CF, DePaulo BM (2006) Accuracy of deception judgments. *Personal Soc Psychol Rev* 10:214–234. https://doi.org/10.1207/s15327957pspr1003_2
- Borchers C (2016) Why the New York Times decided it is now okay to call Donald Trump a liar. *Washington Post*. <https://www.washingtonpost.com/news/the-fix/wp/2016/09/22/why-the-new-york-times-decided-it-is-now-okay-to-call-donald-trump-a-liar/>
- Borchers C (2017) Alex Jones should not be taken seriously, according to Alex Jones's lawyers. *Washington Post*. <https://www.washingtonpost.com/news/the-fix/wp/2017/04/17/trump-called-alex-jones-amazing-joness-own-lawyer-calls-him-a-performance-artist/>
- Braun MT, Swol LMV, Vang L (2015) His lips are moving: Pinocchio effect and other lexical indicators of political deceptions. *Discourse Process* 52:1–20. <https://doi.org/10.1080/0163853X.2014.942833>
- Brown TR (2022) Demystifying mindreading for the law. *Wisconsin Law Review Forward*, 1–11
- Brulle RJ (2018) The climate lobby: A sectoral analysis of lobbying spending on climate change in the USA, 2000 to 2016. *Climatic Change*. <https://doi.org/10.1007/s10584-018-2241-z>
- Brumbach K (2023) Giuliani concedes he made public comments falsely claiming Georgia election workers committed fraud. *Associated Press*. <https://apnews.com/article/giuliani-georgia-election-workers-lawsuit-false-statements-afc64a565ee778c6914a1a69dc756064>
- Bursztyn L, Egorov G, Enikolopov R, Petrova M (2019) Social media and xenophobia: Evidence from Russia (tech. rep.). National Bureau of Economic Research. <https://doi.org/10.3386/w26567>
- Bursztyn L, Rao A, Roth C, Yanagizawa-Drott D (2020) Misinformation during a pandemic (tech. rep.). National Bureau of Economic Research. <https://doi.org/10.3386/w27417>
- Canon DT, Sherman O (2021) Debunking the “Big Lie”: Election Administration in the 2020 Presidential Election. *Pres Stud Q* 51:546–581. <https://doi.org/10.1111/psq.12721>
- Carrasco-Farré C (2022) The fingerprints of misinformation: How deceptive content differs from reliable sources in terms of cognitive effort and appeal to emotions. *Hum Soc Sci Commun* 9:1–18. <https://doi.org/10.1057/s41599-022-01174-9>
- Cataldo JK, Bero LA, Malone RE (2010) “A delicate diplomatic situation”: Tobacco industry efforts to gain control of the Framingham study. *J Clin Epidemiol* 63:841–853. <https://doi.org/10.1016/j.jclinepi.2010.01.021>
- Ceccarelli L (2011) Manufactured scientific controversy: Science, rhetoric, and public debate. *Rhetor Public Aff* 14:195–228
- Constâncio AS, Tsunoda DF, Silva HDFN, Silveira JMD, Carvalho DR (2023) Deception detection with machine learning: a systematic review and statistical analysis. *PLoS One* 18:e0281323. <https://doi.org/10.1371/journal.pone.0281323>
- Cook J, Ellerton P, Kinkead D (2018) Deconstructing climate misinformation to identify reasoning errors. *Environ Res Lett* 13:024018
- Cooley A, Nexon DH (2022) The real crisis of global order: Illiberalism on the rise. *Foreign Aff* 101:103–118
- Darcy O (2021) Fox has quietly implemented its own version of a vaccine passport while its top personalities attack them. *CNN*. <https://edition.cnn.com/2021/07/19/media/fox-vaccine-passport/index.html>
- Davis D, Sinnreich A (2020) Beyond fact-checking: Lexical patterns as lie detectors in Donald Trump's tweets. *Int J Commun* 14:5237–5260
- de Freitas Melo P, Vieira CC, Garimella K, de Melo POV, Benevenuto F (2019) Can WhatsApp counter misinformation by limiting message forwarding? International Conference on Complex Networks and Their Applications, 372–384. https://doi.org/10.1007/978-3-030-36687-2_31
- DellaVigna S, Kaplan E (2007) The fox news effect: media bias and voting. *Q J Econ* 122:1187–1234
- Desikan A, MacKinney T, Kalman C, Carter JM, Reed G, Goldman GT (2023) An equity and environmental justice assessment of anti-science actions during the Trump administration. *J Public Health Policy* 44:147–162. <https://doi.org/10.1057/s41271-022-00390-6>
- Dieterich W, Mendoza C, Brennan T (2016) COMPAS risk scales: Demonstrating accuracy equity and predictive parity. (tech. rep.). Northpoint, Inc
- Dixit P, Mac R (2018) How WhatsApp destroyed a village. *BuzzFeed News*. <https://www.buzzfeednews.com/article/pranavdixit/whatsapp-destroyed-village-lynchings-rainpada-india>
- Douglas A (2018) Disentangling perjury and lying. *Yale J Law Hum* 29:339–374
- Dourado T, Salgado S (2021) Disinformation in the Brazilian pre-election context: Probing the content, spread and implications of fake news about Lula da Silva. *Commun Rev* 24:297–319. <https://doi.org/10.1080/10714421.2021.1981705>
- Dressel J, Farid H (2018) The accuracy, fairness, and limits of predicting recidivism. *Sci Adv* 4:eaa05580
- Ecker UKH, Lewandowsky S, Apai J (2011) Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information. *Q J Exp Psychol* 64:283–310. <https://doi.org/10.1080/17470218.2010.497927>
- Ecker UKH, Lewandowsky S, Cook J, Schmid P, Fazio LK, Brashier N, Kendeou P, Vraga EK, Amazeen MA (2022) The psychological drivers of misinformation belief and its resistance to correction. *Nat Rev Psychol* 1:13–29. <https://doi.org/10.1038/s44159-021-00006-y>
- Eggers AC, Garro H, Grimmer J (2021) No evidence for systematic voter fraud: a guide to statistical claims about the 2020 election. *Proc Natl Acad Sci USA* 118:e2103619118. <https://doi.org/10.1073/pnas.2103619118>
- Enders A, Farhart C, Miller J, Uscinski J, Saunders K, Drochon H (2022) Are republicans and conservatives more likely to believe conspiracy theories? *Polit Behav* 1–24. <https://doi.org/10.1007/s11109-022-09812-3>
- Enders AM, Uscinski JE (2021) Are misinformation, antiscientific claims, and conspiracy theories for political extremists? *Group Processes & Intergroup Relations*
- Fallin A, Grana R, Glantz SA (2013) ‘to quarterback behind the scenes, third-party efforts’: the tobacco industry and the tea party. *Tob Control* 0:1–10. <https://doi.org/10.1136/tobaccocontrol-2012-050815>
- Farber HJ, Neptune ER, Ewart GW (2018) Corrective statements from the tobacco industry: more evidence for why we need effective tobacco control. *Ann Am Thorac Soc* 15:127–130. <https://doi.org/10.1513/annalsats.201711-845gh>
- Farrell H, Schneider B (2018) Common-knowledge attacks on democracy (tech. rep.). Berkman Klein Center for Internet & Society
- Farrell J (2016) Network structure and influence of the climate change counter-movement. *Nat Clim Change* 6:370–374. <https://doi.org/10.1038/nclimate2875>
- Fausset R, Hakim D (2023) Sidney Powell pleads guilty in Georgia Trump case. *New York Times*. <https://www.nytimes.com/2023/10/19/us/sidney-powell-guilty-plea-trump-georgia.html>
- Fazio LK, Brashier NM, Payne BK, Marsh EJ (2015) Knowledge does not protect against illusory truth. *J Exp Psychol General*. <https://doi.org/10.1037/xge0000098>
- Fazio L (2020) Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harv Kennedy School Misinform Rev*. <https://doi.org/10.37016/mr-2020-009>
- Feldman L, Maibach EW, Roser-Renouf C, Leiserowitz A (2012) Climate on cable: the nature and impact of global warming coverage on Fox News, CNN, and MSNBC. *Int J Press/Polit* 17:3–31
- Field H, Vanian J (2023) Tech layoffs ravage the teams that fight online misinformation and hate speech. *CNBC*. <https://www.cnbc.com/2023/05/26/tech-companies-are-laying-off-their-ethics-and-safety-teams-.html>
- Fong A, Roozenbeek J, Goldwert D, Rathje S, van der Linden S (2021) The language of conspiracy: a psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Process Intergroup Relat* 24:606–623. <https://doi.org/10.1177/1368430220987596>
- Francey N, Chapman S (2000) “operation Berkshire”: the international tobacco companies’ conspiracy. *Br Med J* 321:371–374. <https://doi.org/10.1136/bmj.321.7257.371>
- Garrett RK, Bond RM (2021) Conservatives’ susceptibility to political misperceptions. *Sci Adv* 7(23):eabf1234. <https://doi.org/10.1126/sciadv.abf1234>
- Ghanem B, Rosso P, Rangel F (2020) An emotional analysis of false information in social media and news articles. *ACM Trans Internet Technol* 20:19:1–19:18. <https://doi.org/10.1145/3381750>
- Goldberg B (2023) Defanging disinformation’s threat to Ukrainian refugees. *Jigsaw*. <https://medium.com/jigsaw/defanging-disinformations-threat-to-ukrainian-refugees-b164dbbc1c60>
- González-Bailón S, Lazer D, Barberá P, Zhang M, Allcott H, Brown T, Crespo-Tenorio A, Freelon D, Gentzkow M, Guess AM, Iyengar S, Kim YM, Malhotra N, Moehler D, Nyhan B, Pan J, Rivera CV, Settle J, Thorson E, Tucker

- JA (2023) Asymmetric ideological segregation in exposure to political news on Facebook. *Science* 381:392–398. <https://doi.org/10.1126/science.ade7138>
- Graham MH, Yair O (2023) Expressive responding and trump’s big lie. *Polit Behav.* <https://doi.org/10.1007/s11109-023-09875-w>
- Greene KT (2024) Partisan differences in the sharing of low-quality news sources by U.S. political elites. *Polit Commun* 1–20. <https://doi.org/10.1080/10584609.2024.2306214>
- Grice HP (1975) Logic and conversation. In P Cole & JL Morgan (Eds.), *Syntax and semantics*, vol. 3: Speech acts (pp. 41–58). Academic Press
- Grimberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D (2019) Fake news on Twitter during the 2016 U.S. presidential election. *Science* 363:374–378. <https://doi.org/10.1126/science.aau2706>
- Grofman B, Cervas J (2023) Statistical fallacies in claims about ‘massive and widespread fraud’ in the 2020 presidential election: examining claims based on aggregate election results 1,2. *Stat Public Policy* 1–36. <https://doi.org/10.1080/2330443X.2023.2289529>
- Guess AM, Nyhan B, Reifler J (2020a) Exposure to untrustworthy websites in the 2016 U.S. election. *Nat Hum Behav* 4:472–480. <https://doi.org/10.1038/s41562-020-0833-x>
- Guess AM, Lockett D, Lyons B, Montgomery JM, Nyhan B, Reifler J (2020b) “Fake news” may have limited effects on political participation beyond increasing beliefs in false claims. *Harv Kennedy School Misinform Rev* 1(1). <https://doi.org/10.37016/mr-2020-004>
- Guess AM, Nagler J, Tucker J (2019) Less than you think: prevalence and predictors of fake news dissemination on Facebook. *Sci Adv* 5:eau4586. <https://doi.org/10.1126/sciadv.aau4586>
- Harris KR (2022) Real fakes: The epistemology of online misinformation. *Philos Technol* 35. <https://doi.org/10.1007/s13347-022-00581-9>
- Henricksen W, Betz B (2023) The stolen election lie and the freedom of speech. *Penn State Law Review.* <https://doi.org/10.2139/ssrn.4354211>
- Hotez P (2023) Anti-science conspiracies pose new threats to US biomedicine in 2023. *FASEB BioAdvances.* <https://doi.org/10.1096/fba.2023-00032>
- Hruschka TMJ, Appel M (2023) Learning about informal fallacies and the detection of fake news: An experimental intervention. *PLoS One* 18:e0283238
- Hsu SS, Weiner R (2023) Defamed Georgia poll workers who won \$148M from Giuliani sue him again. *Washington Post.* <https://www.washingtonpost.com/dc-md-va/2023/12/18/giuliani-defamation-lawsuit-georgia/>
- Hurley L (2023) Supreme Court blocks restrictions on Biden administration efforts to get platforms to remove social media posts. *NBC News.* <https://www.nbcnews.com/politics/supreme-court/supreme-court-blocks-biden-social-media-curbs-rcna105785>
- Huszár F, Ktena SI, O’Brien C, Belli L, Schlaikjer A, Hardt M (2022) Algorithmic amplification of politics on Twitter. *Proc Natl Acad Sci* 119:e2025334119. <https://doi.org/10.1073/pnas.2025334119>
- Jacobson GC (2021) Donald Trump’s big lie and the future of the republican party. *Pres Stud Q* 51:273–289. <https://doi.org/10.1111/psq.12716>
- Jacobson GC (2023) The dimensions, origins, and consequences of belief in Donald Trump’s Big Lie. *Polit Sci Q* 138:133–166. <https://doi.org/10.1093/psquar/qqac030>
- Jalli N, Idris I (2019) Fake news and elections in two Southeast Asian nations: A comparative study of Malaysia general election 2018 and Indonesia presidential election 2019. *Proceedings of the International Conference of Democratization in Southeast Asia (ICDeSA 2019).* <https://doi.org/10.2991/icdesa-19.2019.30>
- Jung Y, Lee S (2023) Trump vs. the GOP: Political Determinants of COVID-19 Vaccination. *Polit Behav.* <https://doi.org/10.1007/s11109-023-09882-x>
- Kellow CL, Steeves HL (1998) The role of radio in the Rwandan genocide. <https://doi.org/10.1111/j.1460-2466.1998.tb02762.x>
- Kinser S (2020) Science in an age of scrutiny: How scientists can respond to criticism and personal attacks. *Union of Concerned Scientists.* <https://www.ucsusa.org/sites/default/files/2020-09/science-in-an-age-of-scrutiny-2020.pdf>
- Kozyreva A, Herzog SM, Lewandowsky S, Hertwig R, Lorenz-Spreen P, Leiser M, Reifler J (2023) Resolving content moderation dilemmas between free speech and harmful misinformation. *Proc Natl Acad Sci USA* 120:e2210666120. <https://doi.org/10.1073/pnas.2210666120>
- Kozyreva A, Lorenz-Spreen P, Herzog SM, Ecker UKH, Lewandowsky S, Hertwig R, Ali A, Bak-Coleman JB, Barzilai S, Basol M, Berinsky A, Betsch C, Cook J, Fazio LK, Geers M, Guess AM, Huang H, Larreguy H, Maertens R, ... Wineburg S (2024) Toolbox of interventions against online misinformation. *Nat Hum Behav.* <https://doi.org/10.31234/osf.io/x8ejt>
- Kozyreva A, Smillie L, Lewandowsky S (2023) Incorporating psychological science into policy making. *Eur Psychol* 28:206–224. <https://doi.org/10.1027/1016-9040/a000493>
- Kozyreva A, Wineburg S, Lewandowsky S, Hertwig R (2023) Critical ignoring as a core competence for digital citizens. *Curr Dir Psychol Sci* 32:81–88. <https://doi.org/10.1177/09637214221121570>
- Kuklinski JH, Quirk PJ, Schwieder DW, Rich RF (1998) “Just the facts, ma’am”: political facts and public opinion. *Ann Am Acad Political Soc Sci* 560:143–154. <https://doi.org/10.1177/0002716298560001011>
- Kull S, Ramsay C, Lewis E (2003) Misperceptions, the media, and the Iraq war. *Political Sci Q* 118:569–598
- Kumari R, Ashok N, Ghosal T, Ekbal A (2022) *What the fake?* Probing misinformation detection standing on the shoulder of novelty and emotion. *Inf Process Manag* 59:102740. <https://doi.org/10.1016/j.ipm.2021.102740>
- Lackey J (2013) Lies and deception: an unhappy divorce. *Analysis.* <https://doi.org/10.1093/analys/ant006>
- Lagioia F, Rovatti R, Sartor G (2023) Algorithmic fairness through group parities? The case of COMPAS-SAPMOC. *AI & SOCIETY*, 38, 459–478. <https://doi.org/10.1007/s00146-022-01441-y>
- Landman A, Glantz SA (2009) Tobacco industry efforts to undermine policy-relevant research. *Am J Public Health* 99:45–58. <https://doi.org/10.2105/AJPH.2004.050963>
- Lasser J, Aroyehun ST, Simchon A, Carrella F, Garcia D, Lewandowsky S (2022) Social media sharing of low quality news sources by political elites. *PNAS Nexus*, pgacl86. <https://doi.org/10.1093/pnasnexus/pgac186>
- Latour B (2004) Why has critique run out of steam? From matters of fact to matters of concern. *Crit Inq* 30:225–248
- Lebernegg N, Eberl J-M, Tolochko P, Boomgaarden H (2024) Do you speak disinformation? Computational detection of deceptive news-like content using linguistic and stylistic features. *Digit J.* <https://doi.org/10.1080/21670811.2024.2305792>
- Leonhardt D (2021) Red Covid. *New York Times.* <https://www.nytimes.com/2021/09/27/briefing/covid-red-states-vaccinations.html>
- Lerer L (2020) Giuliani in public: ‘it’s a fraud.’ Giuliani in court: ‘This is not a fraud case.’ *New York Times* <https://www.nytimes.com/2020/11/18/us/politics/trump-giuliani-voter-fraud.html>
- Levine S (2023) Angry Fox News chief said fact-checks of Trump’s election lies ‘bad for business’. *The Guardian.* <https://www.theguardian.com/media/2023/mar/29/fox-news-trump-fact-check-election-lies-dominion>
- Lewandowsky S (2020) Willful construction of ignorance: A tale of two ontologies. In R Hertwig & C Engel (Eds.), *Deliberate ignorance: Choosing not to know* (pp. 101–117). MIT Press
- Lewandowsky S, Ballard T, Oberauer K, Benestad R (2016) A blind expert test of contrarian claims about climate data. *Glob Environ Change* 39:91–97. <https://doi.org/10.1016/j.gloenvcha.2016.04.013>
- Lewandowsky S, Ecker UKH, Cook J (2017) Beyond misinformation: understanding and coping with the post-truth era. *J Appl Res Mem Cogn* 6:353–369. <https://doi.org/10.1016/j.jarmac.2017.07.008>
- Lewandowsky S, Kalish ML, Ngang S (2002) Simplified learning in complex situations: Knowledge partitioning in function learning. *J Exp Psychol Gen* 131:163–193. <https://doi.org/10.1037/0096-3445.131.2.163>
- Lewandowsky S, Robertson RE, DiResta R (2023a) Challenges in understanding human-algorithm entanglement during online information consumption. *Perspect Psychol Sci.* <https://doi.org/10.1177/17456916231180809>
- Lewandowsky S, Stritzke WGK, Freund AM, Oberauer K, Krueger JI (2013) Misinformation, disinformation, and violent conflict: From Iraq and the “War on Terror” to future threats to peace. *Am Psychol* 68:487–501. <https://doi.org/10.1037/a0034515>
- Lewandowsky S (2022) Fake news and participatory propaganda. In R Pohl (Ed.), *Cogn illusions* (pp. 324–340). Routledge <https://doi.org/10.4324/9781003154730-23>
- Lewandowsky S, Ecker UKH, Cook J, van der Linden S, Roozenbeek J, Oreskes N (2023b) Misinformation and the epistemic integrity of democracy. *Curr Opin Psychol* 101711. <https://doi.org/10.1016/j.copsyc.2023.101711>
- Lewandowsky S, Pomerantsev P (2022) Technology and democracy: a paradox wrapped in a contradiction inside an irony. *Memory Mind Media* 1. <https://doi.org/10.1017/mem.2021.7>
- Lewandowsky S, van der Linden S (2021) Countering misinformation and fake news through inoculation and prebunking. *Eur Rev Soc Psychol* 32:348–384. <https://doi.org/10.1080/10463283.2021.1876983>
- Li D (2004) Echoes of violence: considerations on radio and genocide in Rwanda. *J Genocide Res* 6:9–27. <https://doi.org/10.1080/1462352042000194683>
- Lin H, Lasser J, Lewandowsky S, Cole R, Gully A, Rand DG, Pennycook G (2023) High level of correspondence across different news domain quality rating sets. *PNAS Nexus* 2:pgad286. <https://doi.org/10.1093/pnasnexus/pgad286>
- Lorenz-Spreen P, Oswald L, Lewandowsky S, Hertwig R (2022) A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nat Hum Behav* 1–28. <https://doi.org/10.1038/s41562-022-01460-1>
- Lu C, Hu B, Li Q, Bi C, Ju X-D (2023) Psychological inoculation for credibility assessment, sharing intention, and discernment of misinformation: systematic review and meta-analysis. *J Med Internet Res* 25:e49255. <https://doi.org/10.2196/49255>

- Martel C, Allen J, Pennycook G, Rand DG (2024) Crowds can effectively identify misinformation at scale. *Perspect Psychol Sci* 19:477–488. <https://doi.org/10.1177/17456916231190388>
- Martel C, Pennycook G, Rand DG (2020) Reliance on emotion promotes belief in fake news. *Cogn Res Princ Implic* 5:47. <https://doi.org/10.1186/s41235-020-00252-3>
- Mattes K, Popova V, Evans JR (2023) Deception detection in politics: can voters tell when politicians are lying. *Polit Behav* 45:395–418. <https://doi.org/10.1007/s11109-021-09747-1>
- McGraw KM (1998) Manipulating public opinion with moral justification. *Ann Am Acad Political Soc Sci* 560:129–142. <https://doi.org/10.1177/0002716298560001010>
- McIntyre L (2018) *Post-truth*. MIT Press
- McLaughlin T (2023) Tail risks for 2024: Prospects for a violent constitutional crisis in the United States (tech. rep. No. 28). Network for Strategic Analysis, Queen's University, Canada
- Mounk Y (2023) *The identity trap*. Penguin Random House
- Müller K, Schwarz C (2021) Fanning the flames of hate: social media and hate crime. *J Eur Econ Assoc* 19:2131–2167. <https://doi.org/10.1093/jeaa/jvaa045>
- Musi E, Aloumpi M, Carmi E, Yates S, O'Halloran K (2022) Developing fake news immunity: Fallacies as misinformation triggers during the pandemic. *Online J Commun Media Technol* 12:e202217. <https://doi.org/10.30935/ojcm/12083>
- Musi E, Reed C (2022) From fallacies to semi-fake news: improving the identification of misinformation triggers across digital media. *Discourse Soc* 33:349–370. <https://doi.org/10.1177/09579265221076609>
- Muzaffar M (2021) Tucker Carlson admits he lies on his show: 'I really try not to... [but] I certainly do'. *The Independent*. <https://www.independent.co.uk/news/world/americas/tucker-carlson-fox-news-dave-rubin-b1919738.html>
- Nadarevic L, Reber R, Helmecke AJ, Köse D (2020) Perceived truth of statements and simulated social media postings: An experimental investigation of source credibility, repeated exposure, and presentation format. *Cogn Res Princ Implic* 5. <https://doi.org/10.1186/s41235-020-00251-4>
- Nan X, Wang Y, Thier K (2022) Why do people believe health misinformation and who is at risk? A systematic review of individual differences in susceptibility to health misinformation. *Soc Sci Med* 314:115398. <https://doi.org/10.1016/j.socscimed.2022.115398>
- Neff A, Fredrickson C (2023) Trump's lawyers face sanctions, discipline, and indictment – how should the legal profession respond? *Just Security*. <https://www.justsecurity.org/90509/trumps-lawyers-face-sanctions-discipline-and-indictment-how-should-the-legal-profession-respond/>
- Neo R (2022) A cudgel of repression: analysing state instrumentalisation of the 'fake news' label in Southeast Asia. *Journalism* 23:1919–1938. <https://doi.org/10.1177/1464884920984060>
- Nieminen S, Sankari V (2021) Checking PolitiFact's fact-checks. *J Stud* 22:358–378. <https://doi.org/10.1080/1461670x.2021.1873818>
- Nix N, Menn J (2023) These academics studied falsehoods spread by Trump. Now the GOP wants answers. *Washington Post*. <https://www.washingtonpost.com/technology/2023/06/06/disinformation-researchers-congress-jim-jordan/>
- Nix N, Zakrzewski C, Menn J (2023) Misinformation research is buckling under GOP legal attacks. *Washington Post*. <https://www.washingtonpost.com/technology/2023/09/23/online-misinformation-jim-jordan/>
- Nyberg D (2023) The passive revolution is televised: The dominant ideology of media capitalism. *Organization*, 13505084231180288. <https://doi.org/10.1177/13505084231180288>
- Ognyanova K, Lazer D, Robertson RE, Wilson C (2020) Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harv Kennedy School Misinform Rev*. <https://doi.org/10.37016/mr-2020-024>
- Oreskes N, Conway EM (2010) *Merchants of doubt*. Bloomsbury Publishing
- Oreskes N, & Conway EM (2023) *The big myth*. New York City: Bloomsbury Publishing
- O'Toole F (2022) We don't know ourselves: A personal history of Ireland since 1958. *Head of Zeus*
- Painter DL, Fernandes J (2022) "the big lie": How fact checking influences support for insurrection. *Am Behav Sci* 000276422211031. <https://doi.org/10.1177/00027642221103179>
- Papantoniou K, Papadakis P, Patkos T, Flouris G, Androutsopoulos I, Plexousakis D (2021) Deception detection in text and its relation to the cultural dimension of individualism/collectivism. *Nat Lang Eng* 28:545–606. <https://doi.org/10.1017/s1531324921000152>
- Peltz M (2023) New details in Dominion suit reveal damning evidence of deception in Fox News' 2020 election coverage. *Mediamatters*. <https://www.mediamatters.org/foxdominion-lawsuit/new-details-dominion-suit-reveal-damning-evidence-deception-fox-news-2020>
- Peng W, Lim S, Meng J (2023) Persuasive strategies in online health misinformation: a systematic review. *Inf Commun Soc* 26:2131–2148. <https://doi.org/10.1080/1369118X.2022.2085615>
- Pennycook G, Rand DG (2021) Research note: Examining false beliefs about voter fraud in the wake of the 2020 presidential election. *Harvard Kennedy School (HKS) Misinform Rev* 2. <https://doi.org/10.37016/mr-2020-51>
- Pennycook G, Rand DG (2019) Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.1806781116>
- Pereira FB, Bueno NS, Nunes F, Pavão N (2022) Fake news, fact checking, and partisanship: the resilience of rumors in the 2018 Brazilian elections. *J Polit* 84:2188–2201. <https://doi.org/10.1086/719419>
- Pereira FB, Bueno NS, Nunes F, Pavão N (2023) Inoculation reduces misinformation: experimental evidence from multidimensional interventions in Brazil. *J Exp Polit Sci* 1–12. <https://doi.org/10.1017/xps.2023.11>
- Persad G, Emanuel EJ, Sangeon S, Glickman A, Phillips S, Largent EA (2021) Public perspectives on COVID-19 vaccine prioritization. *JAMA Netw Open* 4:e217943. <https://doi.org/10.1001/jamanetworkopen.2021.7943>
- Pillai RM, Fazio LK (2023) Explaining why headlines are true or false reduces intentions to share false information. *Collabra: Psychol* 9. <https://doi.org/10.1525/collabra.87617>
- Pinna M, Picard L, Goessmann C (2022) Cable news and COVID-19 vaccine uptake. *Sci Rep* 12:16804. <https://doi.org/10.1038/s41598-022-20350-0>
- Polantz K (2021) Lawyers sanctioned for 'conspiracy theory' election fraud lawsuit. *CNN*. <https://edition.cnn.com/2021/08/04/politics/lawyers-colorado-20-election/index.html>
- Porter E, Wood TJ (2021) The global effectiveness of fact-checking: evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *Proc Natl Acad Sci USA* 118:e2104235118. <https://doi.org/10.1073/pnas.2104235118>
- Pothos EM, Lewandowsky S, Basieva I, Barque-Duran A, Tapper K, Khrennikov A (2021) Information overload for (bounded) rational agents. *Proc R Soc B Biol Sci* 288:20202957. <https://doi.org/10.1098/rspb.2020.2957>
- Prike T, Butler LH, Ecker UKH (2024) Source-credibility information and social norms improve truth discernment and reduce engagement with misinformation online. *Sci Rep* 14:6900. <https://doi.org/10.1038/s41598-024-57560-7>
- Proctor RN (2011) *Golden holocaust: Origins of the cigarette catastrophe and the case for abolition*. University of California Press
- Proctor RN (2012) The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll. *Tob Control* 21(2):87–91. <https://doi.org/10.1136/tobaccocontrol-2011-050338>
- Pröllochs N, Bär D, Feuerriegel S (2021) Emotions explain differences in the diffusion of true vs. false social media rumors. *Sci Rep* 11:22721. <https://doi.org/10.1038/s41598-021-01813-2>
- Reid T (2022) Voting machine problems in Arizona seized on by Trump, election deniers. *Reuters*. <https://www.reuters.com/world/us/voting-machine-problems-battleground-arizona-seized-by-trump-election-deniers-2022-11-08/>
- Rogers T, Zeckhauser R, Gino F, Norton MI, Schweitzer ME (2017) Artful paltering: the risks and rewards of using truthful statements to mislead others. *J Personal Soc Psychol* 112:456–473
- Rozenbeek J, van der Linden S, Goldberg B, Rathje S, Lewandowsky S (2022) Psychological inoculation improves resilience against misinformation on social media. *Sci Adv* 8:eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- Rozenbeek J, Suiter J, Culloty E (2023) Countering misinformation: evidence, knowledge gaps, and implications of current interventions. *Eur Psychol*. <https://doi.org/10.31234/osf.io/b52um>
- Rumpler E, Feldman JM, Bassett MT, Lipsitch M (2023) Fairness and efficiency considerations in COVID-19 vaccine allocation strategies: A case study comparing front-line workers and 65–74 year olds in the United States. *PLOS Glob Public Health* 3:e0001378. <https://doi.org/10.1371/journal.pgph.0001378>
- Rutenberg J, Myers SL (2024) How Trump's allies are winning the war over disinformation. *New York Times*. <https://www.nytimes.com/2024/03/17/us/politics/trump-disinformation-2024-social-media.html>
- Salek TA (2023) Deflecting deliberation through rhetorical nihilism: "Stop the Steal" as an unethical and intransigent rival public. *Commun Democracy* 57:94–118. <https://doi.org/10.1080/27671127.2023.2202744>
- Scarcella M (2023) US Supreme Court rebuffs lawyers punished after 'woeful' suit backing Trump. *Reuters*. <https://www.reuters.com/legal/us-supreme-court-rebuffs-lawyers-punished-after-woeful-suit-backing-trump-2023-10-02/>
- Shahid F, Vashista A (2023) Decolonizing Content Moderation: Does Uniform Global Community Standard Resemble Utopian Equality or Western Power Hegemony? *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. <https://doi.org/10.1145/3544548.3581538>
- Simonov A, Sacher S, Dubé J-P, Biswas S (2020) The persuasive effect of Fox News: Non-compliance with social distancing during the Covid-19 pandemic (tech. rep.). National Bureau of Economic Research. <https://doi.org/10.3386/w27237>

- Simonov A, Sacher S, Dubé J-P, Biswas S (2022) Frontiers: the persuasive effect of fox news: noncompliance with social distancing during the COVID-19 pandemic. *Mark Sci* 41:230–242. <https://doi.org/10.1287/mksc.2021.1328>
- Smith P, Bansal-Travers M, O'Connor R, Brown A, Banthin C, Guardino-Colket S, Cummings K (2011) Correcting over 50 years of tobacco industry misinformation. *Am J Prev Med* 40:690–698
- Soroka S, Fournier P, Nir L (2019) Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proc Natl Acad Sci USA* 116:18888–18892. <https://doi.org/10.1073/pnas.1908369116>
- Stapleton A (2016) No, you can't vote by text message. CNN. <https://edition.cnn.com/2016/11/07/politics/vote-by-text-message-fake-news/index.html>
- Starbird K, DiResta R, DeButts M (2023) Influence and Improvisation: participatory disinformation during the 2020 US election. *Soc Media Soc* 9:20563051231177943. <https://doi.org/10.1177/20563051231177943>
- Supran G, Rahmstorf S, Oreskes N (2023) Assessing ExxonMobil's global warming projections. *Science* 379(6628):eabk0063. <https://doi.org/10.1126/science.abk0063>
- Supran G, Oreskes N (2017) Assessing ExxonMobil's climate change communications (1977–2014). *Environ Res Lett* 12:084019. <https://doi.org/10.1088/1748-9326/aa815f>
- Supran G, Oreskes N (2021) Rhetoric and frame analysis of ExxonMobil's climate change communications. *One Earth*. <https://doi.org/10.1016/j.oneear.2021.04.014>
- Swire-Thompson B, Ecker UKH, Lewandowsky S, Berinsky AJ (2020) They might be a liar but they're my liar: Source evaluation and the prevalence of misinformation. *Polit Psychol* 41:21–34. <https://doi.org/10.1111/pops.12586>
- Takshid Z (2021) Regulating social media in the global south. *Vanderbilt J Entertain Technol Law* 24:1–56
- Tenove C (2020) Protecting democracy from disinformation: Normative threats and policy responses. *Int J Press/Polit* 25:517–537. <https://doi.org/10.1177/1940161220918740>
- Terkel A, Timm JC, Gregorian D (2023) Here's what fox news was trying to hide in its dominion lawsuit redactions. NBC News. <https://www.nbcnews.com/politics/elections/dominion-releases-previous-redacted-slides-fox-news-lawsuit-rcna77257>
- U.S. House of Representatives Judiciary Committee. (2023). News release: Jim Jordan on why the select subcommittee on the weaponization of the federal government is necessary | House Judiciary Committee Republicans. <http://judiciary.house.gov/media/press-releases/jim-jordan-on-why-the-select-subcommittee-on-the-weaponization-of-the-federal>
- Uscinski JE, Parent JM (2014) *American conspiracy theories*. Oxford University Press
- Uscinski JE (2015) The epistemology of fact checking (is still naive): Rejoinder to Amazeen. *Crit Rev* 27:243–252. <https://doi.org/10.1080/08913811.2015.1055892>
- Van Der Zee S, Poppe R, Havrileck A, Baillon A (2021) A personal model of trumpery: linguistic deception detection in a real-world high-stakes setting. *Psychol Sci*. <https://doi.org/10.1177/09567976211015941>
- van Doorn M (2023) Advancing the debate on the consequences of misinformation: clarifying why it's not (just) about false beliefs. *Inquiry* 0:1–27. <https://doi.org/10.1080/0020174X.2023.2289137>
- Vese D (2022) Governing fake news: the regulation of social media and the right to freedom of expression in the era of emergency. *Eur J Risk Regul* 13:477–513. <https://doi.org/10.1017/err.2021.48>
- Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359:1146–1151. <https://doi.org/10.1126/science.aap9559>
- Walasek L, Brown GDA (2023) Incomparability and incommensurability in choice: No common currency of value? *Perspect Psychol Sci*. <https://doi.org/10.1177/17456916231192828>
- Wallace J, Goldsmith-Pinkham P, Schwartz JL (2023) Excess death rates for republican and democratic registered voters in Florida and Ohio during the COVID-19 pandemic. *JAMA Internal Med*. <https://doi.org/10.1001/jamainternmed.2023.1154>
- Wanless A, Berk M (2019) The audience is the amplifier: Participatory propaganda. In P Baines, N O'Shaughnessy, & N Snow (Eds.), *The sage handbook of propaganda* (pp. 85–104). Sage
- Wardle C, Derakhshan H (2017) Information disorder: Toward an interdisciplinary framework for research and policymaking (tech. rep.). Council of Europe. <https://rm.coe.int/information-disorder-report-version-august-2018/16808c9c77>
- Weinschenk AC, Panagopoulos C, van der Linden S (2021) Democratic norms, social projection, and false consensus in the 2020 U.S. presidential election. *J Polit Mark* 20:255–268. <https://doi.org/10.1080/15377857.2021.1939568>
- West D (2023) We shouldn't turn disinformation into a constitutional right. Brookings Institution. <https://www.brookings.edu/articles/we-shouldnt-turn-disinformation-into-a-constitutional-right/>
- Williams D (2021) Motivated ignorance, rationality, and democratic politics. *Synthese* 198:7807–7827. <https://doi.org/10.1007/s11229-020-02549-8>
- Williams D (2022) The marketplace of rationalizations. *Econ Philosophy*, 1–25. <https://doi.org/10.1017/S0266267121000389>
- Wu T (2017) *The attention merchants*. Atlantic Books
- Yee AK (2023a) Information deprivation and democratic engagement. *Philos Sci* 90:1110–1119. <https://doi.org/10.1017/psa.2023.9>
- Yee AK (2023b) Machine learning, misinformation, and citizen science. *Eur J Philosophy* 13. <https://doi.org/10.1007/s13194-023-00558-1>
- Zakrzewski C, Lima C, Harwell D (2023) What the Jan. 6 probe found out about social media, but didn't report. *Washington Post*. <https://www.washingtonpost.com/technology/2023/01/17/jan6-committee-report-social-media/>

Acknowledgements

SL acknowledges financial support from the European Research Council (ERC Advanced Grant 101020961 PRODEMINFO), the Humboldt Foundation through a research award, the Volkswagen Foundation (grant "Reclaiming individual autonomy and democratic discourse online: How to rebalance human and algorithmic decision making"), and the European Commission (Horizon 2020 grants 964728 JITSUVAX and 101094752 SoMe4Dem). SL also receives funding from Jigsaw (a technology incubator created by Google) and from UK Research and Innovation through EU Horizon replacement funding grant number 10049415. UKHE acknowledges support from the Australian Research Council (grant FT190100708). For the purpose of open access, the author(s) has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

The first author created the first draft and all other authors contributed additional material and comments and suggestions and participated jointly in the editing and revision process.

Competing interests

SL, JR, and SvDL have received funding from Google Jigsaw for empirical work on inoculation against misinformation and continue to collaborate with Jigsaw. NO has received funding from the Rockefeller Family Fund to support research on fossil fuel industry disinformation. She has also served as a consultant to the law firm Sher-Edling, who are representing several counties in California suing the fossil fuel industry, and as an expert witness in the defamation case of climate scientist Michael Mann. The remaining authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors that would require consent.

Additional information

Correspondence and requests for materials should be addressed to Stephan Lewandowsky.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024