



Lewandowsky, S., & Pomerantsev, P. (2021). Technology and democracy: a paradox wrapped in a contradiction inside an irony. *Memory, Mind & Media (MMM)*, 1, [e5].
<https://doi.org/10.1017/mem.2021.7>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1017/mem.2021.7](https://doi.org/10.1017/mem.2021.7)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Cambridge University Press at [10.1017/mem.2021.7](https://doi.org/10.1017/mem.2021.7). Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

RESEARCH ARTICLE

Technology and democracy: a paradox wrapped in a contradiction inside an irony

Stephan Lewandowsky¹  and Peter Pomerantsev²

¹School of Psychological Science, University of Bristol, Bristol BS8 1TU, UK and ²SNF Agora Institute, Johns Hopkins University, Baltimore, Maryland

Corresponding author: Stephan Lewandowsky, email: stephan.lewandowsky@bristol.ac.uk

(Received 5 November 2021; accepted 8 November 2021)

Abstract

Democracy is in retreat around the globe. Many commentators have blamed the Internet for this development, whereas others have celebrated the Internet as a tool for liberation, with each opinion being buttressed by supporting evidence. We try to resolve this paradox by reviewing some of the pressure points that arise between human cognition and the online information architecture, and their fallout for the well-being of democracy. We focus on the role of the attention economy, which has monetised dwell time on platforms, and the role of algorithms that satisfy users' presumed preferences. We further note the inherent asymmetry in power between platforms and users that arises from these pressure points, and we conclude by sketching out the principles of a new Internet with democratic credentials.

Keywords: Democracy; populism; misinformation; social media; Internet

The mission of *Memory, Mind & Media* is to document and explore the impact of media and technology on human, social and cultural remembering and forgetting. In this article, we set out the key challenges for the field, and hence the core issues and ideas for the journal, through the lens of the unique cognitive pressure points that create tension between the online information ecology and democratic discourse and governance.

Numerous indicators suggest that democracy is in retreat globally (Freedom House 2020; Lührmann and Lindberg 2020). Even countries that had been considered stable democracies, such as the United States (US) and the United Kingdom (UK), have recently witnessed events that are incompatible with democratic governance and the rule of law, such as the armed assault on the U.S. Capitol in 2021 and the unlawful suspension of the British parliament in 2019.

Although the symptoms and causes of democratic backsliding are complex and difficult to disentangle, the Internet and social media are frequently blamed in this context. For example, social media has been identified as a tool of autocrats (Deibert 2019).

Empirical support for this assertion arises from the finding that the more committed autocratic regimes are to prevent an independent public sphere, the more likely they are to introduce the Internet (Rød and Weidmann 2015). In Western democracies, recent evidence suggests that social media can *cause* some anti-democratic political behaviours

ranging from ethnic hate crimes to voting for populist parties (Bursztyn et al 2019; Müller and Schwarz 2019; Allcott et al 2020; Schaub and Morisi 2020). Social media have also been blamed for increasing political polarisation (Van Bavel et al 2021). Some scholars have openly questioned whether democracy can survive the Internet (Persily 2017).

In the opposing corner, social media has been heralded as ‘liberation technology’ (Tucker et al 2017), owing to its role in the ‘Arab Spring’, the Iranian Green Wave Movement of 2009, and other instances in which it mobilised the public against autocratic regimes. Similarly, protest movements in the US, Spain, Turkey, and Ukraine rely on social media platforms for the coordination of collective action and to transmit emotional and motivational messages (Jost et al 2018). A recent field experiment in an ethnically highly polarised society, Bosnia and Herzegovina, found that people who continued to use Facebook reported greater outgroup regard than a group that voluntarily deactivated Facebook for the same time period (Asimovic et al 2021).

The fundamental paradox

This is the fundamental paradox of the Internet and social media: They erode democracy and they expand democracy. They are the tools of autocrats and they are the tools of activists. They make people obey and they make them protest. They provide a voice to the marginalised and they give reach to fanatics and extremists. And all of these conflicting views are seemingly supported by analysis or empirical evidence, rendering resolution of this paradox difficult.

We have proposed elsewhere that to understand this basic paradox, we must examine the unique pressure points that arise when human cognition is let loose on the Internet (Kozyreva et al 2020; Lewandowsky et al 2020; Lorenz-Spreen et al 2020). The interaction between fundamental human cognitive attributes and the architecture of the information ecology have created a perfect storm for democracy. Here, we focus on a subset of these pressure points and highlight how they, in turn, also contain intrinsic ironies and paradoxes.

The attention economy

Our attention has been commodified (Wu 2017). When we use a ‘free’ product online, we are the product. The more time we spend watching YouTube videos or checking our Facebook newsfeed, the more advertising revenue is generated for the platforms. This commodification of attention is an inescapable driver of online behaviour that has several contradictory consequences. On the positive side, the fact that dwell time online has become revenue-generating currency has enabled the creation of a vast array of – seemingly – free services. YouTube is free to use and provides nearly unlimited entertainment options. Google offers a suite of tools beyond its search engine, from email to document creation, that support countless endeavours free of charge. Facebook permits us to stay in touch with friends and family, and we can use WhatsApp to make video calls with people all around the world at no cost. The array of free services available online is impressive by any measure.

But those free services are not truly free – on the contrary, they incur considerable costs that are often external to the interactions we intentionally engage in. One implication of the conversion of dwell time into revenue-generating currency is that the platforms will naturally try to present us with captivating information to retain our attention. This commercial incentive structure is potentially problematic because people are known to attend to news that is predominantly negative (Soroka et al 2019) or awe-inspiring (Berger and Milkman 2012). People also preferentially share messages couched in moral-emotional language (Brady et al 2017). It is unsurprising, therefore, that ‘fake news’ and misinformation have become so prevalent online because false

content – which by definition is freed from factual constraints – can exploit this attentional bias: misinformation on Facebook during the 2016 U.S. presidential campaign was particularly likely to provoke voter outrage (Bakir and McStay 2018) and fake news titles have been found to be substantially more negative in tone, and display more negative emotions such as disgust and anger, than real news titles (Paschen 2019). The flood of disinformation and online outrage is, therefore, arguably a price we pay for the ‘free’ services provided by the platforms.

Although human attentional biases did not suddenly change just because the Internet was invented – the adage that ‘when it bleeds, it leads’ is probably as old as journalism itself – web technology has turbo-charged those biases in at least two ways. First, the sheer quantity of information online has measurable adverse consequences for our ‘collective mind’ and societal memories. Whereas in 2013 the most popular hashtags on Twitter remained popular for 17.5 h, by 2016 a hashtag’s life in the limelight had dropped to 11.9 h (Lorenz-Spreen et al 2019). A similar decline in our collective attention span was observed for Google queries and movie ticket sales (Lorenz-Spreen et al 2019). It is unsurprising that political accountability will become more difficult in societies with a shorter attention span: if a leader’s original transgression is forgotten in a few hours, the public appetite for accountability is unlikely to be lasting (Giroux and Bhattacharya 2016). Even highly consequential events can seemingly disappear without leaving much of a trace: When British Prime Minister Boris Johnson prorogued (ie, shut down) Parliament on 24 August 2019 to escape further scrutiny of his Brexit plans, public interest was initially intense. After this prorogation was found to be unlawful by the Supreme Court on 24 September 2019, public interest in the issue, as measured by Google Trends, dissipated by 93% within 5 days. Within 2 months, public interest in prorogation returned to the near-zero level observed before the prorogation, when hardly anyone in Britain even knew the term ‘prorogation’ existed.¹ Johnson went on to win an election a few months later by a landslide.

The problems arising from a shortened attention span are compounded by the fact that information overload generally makes it harder for people to make good decisions about what to look at, what to spend time on, what to believe, and what to share (Hills et al 2013; Hills 2019). Choosing a paper to purchase at a newsstand requires a single decision. Our Twitter feed or Facebook newsfeed confronts us with the need for a multitude of micro-decisions for every article or post. Although these repeated micro-decisions open the door to greater diversity in our news diet, they also increase the probability that at least some of our chosen sources fail to be trustworthy. Worse yet, information overload can also contribute to polarisation and dysfunctional disagreement between well-meaning and rational actors (Pothos et al 2021). That is, despite their good-faith efforts, overload may prevent actors from forming compatible mental representations of complex problems. Excessive complexity mandates a simplification of representations, and this, in turn, necessarily introduces potential incompatibilities between actors that may result in irresolvable disagreement (Pothos et al 2021).

The second turbocharger of human cognitive biases by online technologies relies on the exact measurement of our responses to information. Facebook has access to our every click while we are on the platform, and Facebook can use that information for continual personalised refinement of our information diet through the platform’s algorithms.

The Jekyll and Hyde of the algorithm

Most of the information we consume online is shaped and curated by algorithms. YouTube, by default, keeps playing videos we are presumed to like based on inferences by its recommender system. Facebook’s newsfeed is curated by a sophisticated algorithm,

and Google's search results are customised according to numerous parameters. Algorithms are an essential tool to harness the abundance of information on the web: Googling 'Georgia' should return different results in Atlanta than in Tbilisi, and without such intelligent filtering, useful information would most likely remain inaccessible. Algorithms can also help us satisfy our preferences, for example, when recommender systems help us find movies, books, or restaurants that we are likely to enjoy (Ricci et al 2015). It is unsurprising, therefore, that the public is mainly appreciative of algorithms and customisation in those contexts (Kozyreva et al 2021).

There are, however, several darker sides to algorithms. The first problem is that algorithms ultimately serve the interests of the platforms rather than the users. An ironic consequence of this is that in the relentless pursuit of increasing dwell time, algorithms may eagerly satisfy our presumed momentary preferences even if that reduces our long-term well-being. In the same way that strategically placed junk food in the supermarket can satisfy our cravings while also propelling an obesity epidemic, algorithms may satisfy our momentary desire for emotional engagement while contributing to the formation of sealed anti-democratic communities (Kaiser and Rauchfleisch 2020). Unconstrained preference satisfaction may ironically create fractionated and polarised societies (Pariser 2011).

The second problem with algorithms is that their design and operation are proprietary and not readily subject to public scrutiny. Most algorithms operate as 'black boxes' where neither individual users nor society, in general, know why search results or social media feeds are curated in a particular way (Pasquale 2015). At present, knowledge about the algorithms can only be obtained by 'reverse engineering' (Diakopoulos 2015), that is, by seeking to infer an algorithm's design based upon its observable behaviour.

Reverse engineering can range from the relatively simple (eg, examining which words are excluded from auto-correct on the iPhone; Keller 2013) to the highly complex (eg, an analysis of how political ads are delivered on Facebook; Ali et al 2019). Reverse engineering has uncovered several problematic aspects of algorithms, such as discriminatory advertising practices and stereotypical representations of Black Americans in Google Search (Sweeney 2013; Noble 2018) and in the autocomplete suggestions that Google provides when entering search terms (Baker and Potts 2013). At the time of this writing, a Facebook whistle-blower revealed further information about how content is being highlighted on the platform. It transpired that any content that made people angry – which was disproportionately likely to include misinformation, toxicity, and low-quality news – was given particular prominence in people's newsfeed. Facebook thus 'systematically amped up some of the worst of its platform, making it more prominent in users' feeds and spreading it to a much wider audience' (Merrill and Oremus 2021).

The opacity of algorithms allows platforms to drench users in information that may be detrimental to democratic health. Even ignoring the specifics of content, algorithmic opacity also contributes to a general imbalance of power between platforms and users that can only be unhealthy in a democracy.

The asymmetry of power

The platforms know much about their users – and even about people who are not on their platforms (Garcia 2017) – and deploy that knowledge to shape our information diets. By contrast, citizens know little about what data the platforms hold and how these data are used (Lorenz-Spreen et al 2020). For example, Facebook 'likes' can be used to infer our personality through machine learning with considerable accuracy (Youyou et al 2015). Knowledge of just a few likes raises machine-learning performance above that of work colleagues, and with knowledge of 300 likes, the performance of the machine exceeds that of one's spouse (Youyou et al 2015). In stark contrast to the power of machine learning, a

substantial share of people does not even know that their Facebook newsfeed is curated based on personal data (Eslami et al 2015; Rader and Gray 2015; Powers 2017), with estimates of this lack of awareness ranging from 27 to 62.5 per cent.

Asymmetry in knowledge translates into an asymmetry of power: To keep others under surveillance while avoiding equal scrutiny oneself is the most important form of authoritarian political power (Balkin 2008; Zuboff 2019). Similarly, to know others while revealing little about oneself is the most important form of commercial power in an attention economy. When Facebook recently shut down the accounts of researchers who were studying how misinformation spreads and how users are targeted on the platform (Edelson and McCoy 2021), it did not do so to preserve users' privacy as it claimed. That claim was quickly and thoroughly rejected by the Federal Trade Commission. Facebook shut down the researchers to preserve its asymmetrical power advantage by preventing an examination of how it operates. It is this power asymmetry that renders the freedom and choice offered by the Internet largely illusory.

The illusion of freedom and choice

Everyone gets a voice on the Internet. On the positive side of the ledger, there is evidence that access to the Internet leads to enhanced transparency and reduction of corruption. In a cross-national analysis of 157 countries, Starke et al (2016) showed that Internet access was associated with a significant reduction in official corruption. On the more negative side of the ledger, a single tweet can trigger a cascade of adverse events. The 'pizzagate' affair of 2016 was triggered by a baseless accusation that the Democratic party was operating a paedophilia ring out of the basement of a pizza parlour in Washington, D.C. This conspiracy theory was eventually picked up by mainstream media, and ultimately an armed individual entered the pizza parlour and fired shots inside in search of a (non-existent) basement (Fisher et al 2016).

The ambivalent consequences of unfettered access to the Internet are amplified by the opportunities offered for manipulation through targeted advertising. All advertising and political speech seek to persuade. Manipulation differs from persuasion by furtively exploiting a target's weaknesses and vulnerabilities to steer their behaviour in a desired direction (Susser et al 2019). The fact that Facebook 'likes' permit inferences about a user's personality (Youyou et al 2015), combined with the fact that advertisers can select target audiences based on those likes (coded as users' interests), offers an opportunity for targeted manipulation on a global scale and without any transparency. Research suggests that single individuals or households can be targeted with messages using Facebook's ad delivery services (Faizullabhoj and Korolova 2018). Although the effectiveness of such 'microtargeting' of messages is subject to debate (eg, Matz et al 2017 vs. Eckles et al 2018), there is no question that targeting of political messages at individuals (or small numbers of individuals) facilitates the dissemination of disinformation because political opponents cannot know what is being said and hence cannot rebut false information (Heawood 2018). Similarly, microtargeting allows politicians to make multiple incompatible promises to different audiences without anyone being able to track and point out those incompatibilities (Heawood 2018). A recent pertinent example arose during the German parliamentary election in September 2021. The Liberal Democratic Party (FDP) was found to target Facebook users with 'green' interests with a message that identified the party with 'more climate protection' through a regulatory upper limit on CO₂ emissions. At the same time, the FDP targeted frequent travellers on Facebook with an ad that promised 'no state intervention or restrictions of freedom or prohibitions' to address climate change.²

Unsurprisingly, the public overwhelmingly rejects this type of manipulative targeting (Kozyreva et al 2021).³

Everyone may get a voice on the Internet. But everyone is also exposed to a cacophony of voices whose origin may be obscured and that may seek to manipulate rather than inform. The power to design and deliver manipulative messages that form our society's collective memory rests with advertisers and platforms rather than citizens. For now at least, the freedom and choice offered by the Internet, therefore, remains largely illusory.

Building a better Internet

Our preceding analysis illustrates the fundamental paradox of the online media environment: On the one hand, there is more information than ever before, but we know less than ever about how that information is produced, targeted, organised, and distributed. Citizens do not know why algorithms show them one thing and not another, or which of their own data are being used to target them and why. Citizens have little way of knowing about the vast social engineering experiments tech companies play with as they fiddle with their algorithms. Citizens do not even know if their basic rights are being infringed by manipulative algorithms and advertisers. We believe that democratic societies would never have consented to any of those consequences of the Internet if they had been known ahead of time or if the Internet had been designed with those attributes in mind. It is only because the Internet evolved, one technological innovation and one tweak to an algorithm at a time, that democracies are only now realising what they are confronting.

What, then, should the online experience be like for a person in a democracy? How can we design and build a better Internet? We have both been involved in developing specific recommendations for a better Internet (eg, Kozyreva et al 2020; Lewandowsky et al 2020; Lorenz-Spreen et al 2020; Applebaum and Pomerantsev 2021). Here, we focus on one aspect only, namely the power asymmetry between platforms and users and how it might be redressed.

In an Internet with democratic credentials, users would be able to understand which of their own data have been used to target them and why. Users would know why algorithms show them one thing and not another. During elections, people would immediately understand how different campaigns target different people with different messages, who is behind campaigns, and how much they spend.

Online anonymity is a basic right. People should be allowed to 'wear a mask' online for reasons of safety and many others. But the receiver of information should also have the right to understand whether they are being targeted by a real person (whether anonymous or not), or by a political campaign, a corporation, or a state that is pretending to be a real person. 'Troll farms', bot nets, and other forms of mass coordinated inauthentic activity should be clearly identified as such.

An empowered online citizen would also have far greater control over their own data and would be able to regulate how others use it. There may be instances where, for example, one might be comfortable with sharing one's data with a national health service.

But there should be strict guardrails that do not allow user data to be used further by data brokers.

And as individuals should have more oversight and control over the information environment all around them, so should the public have greater oversight and control over tech companies in general. The public need to be able to understand what social engineering experiments the companies tinker with, what their impacts are, and how the tech companies track the consequences of these experiments.

Likewise, algorithmic transparency is essential. This does not mean that companies have to reveal their proprietary source code. They do, however, need to explain the purpose of adjustments they make to their algorithms, and the changes these bring about. If

algorithms infringe on people's rights, such as in cases where algorithms produce advertising that disadvantages minorities, the public need to have oversight over what the companies are doing to rectify these discriminatory practices. Such algorithmic transparency needs to be backed up with regulatory teeth: regulators should have the right to spot-check how companies are continually analysing and mitigating negative effects of their own design decisions.

But regulation needs to go beyond just mitigating the bad and setting standards. It needs to encourage 'the good' too. We must design regulations that encourage the development of 'civic tech'; that is, technology that is meant to benefit individuals and strengthen democratic processes. Such technology would be created in the public interest, and not driven by short-term profit motives to extract people's personal data and then sell it on.

As Ethan Zuckerman of the University of Massachusetts argues⁴, we are in a similar place in the development of the Internet as we were with radio at the start of the 20th century. Back in the 1920s, in the UK, Lord Reith fought for the existence of public interest broadcasting to balance the polarising impact of press barons and the rising power of radio-enhanced dictatorships. The result was the creation of the BBC. What would be the online equivalent of that today? We do not know. This illustrates the magnitude of the task ahead. It may be daunting, but that should concern us less than the conflict between current technologies and democracy that is driven, in part, by known limitations of human attention, memory, and cognition. The mission of *Memory, Mind & Media* is precisely aimed at those limitations and conflicts, and the journal is, therefore, poised to make a contribution to what we consider the defining political battle of the 21st century – the battle between technological hegemony and survival of democracy.

Funding. The first author was supported by funding from the Humboldt Foundation in Germany through a research award, and by an ERC Advanced Grant (PRODEMINFO). The preparation of this paper was also facilitated by a grant from the Volkswagen Foundation for the project 'Reclaiming individual autonomy and democratic discourse online'.

Conflict of Interest. The authors declare no competing interests.

Notes

¹ <https://trends.google.com/trends/explore/TIMESERIES/1635669000?hl=en-GB&tz=0&date=today+5-y&geo=GB&q=prorogation&sn=3>.

² <https://targetleaks.de/>.

³ Recent transparency measures (eg, Facebook's 'ad library') are insufficient to analyze parties' expenditure on microtargeting and what content has been shown (Dommert and Power 2019). The ad library is also missing more than 100,000 political ads (Edelson and McCoy 2021). This difficulty is likely to persist because ads on Facebook are delivered by a continually evolving algorithm, known as AdTech, that auctions off ads on a second-to-second basis based on live analysis of user data (Ali et al 2019).

⁴ <https://www.umass.edu/spp/news/case-digital-public-infrastructure>.

References

- Ali M, Sapiezynski P, Korolova A, Mislove A and Rieke A (2019) *Ad delivery algorithms: the hidden arbiters of political messaging*. Tech. Rep. Available at <https://arxiv.org/pdf/1912.04255.pdf>, accessed 19 April 2020.
- Allcott H, Braghieri L, Eichmeyer S and Gentzkow M (2020) The welfare effects of social media. *American Economic Review* **110**, 629–676. doi:10.1257/aer.20190658
- Applebaum A and Pomerantsev P (2021) *How to put out democracy's dumpster fire*. Available at <https://www.theatlantic.com/magazine/archive/2021/04/the-internet-doesnt-have-to-be-awful/618079/>, accessed 4 August 2021.
- Asimovic N, Nagler J, Bonneau R and Tucker JA (2021) Testing the effects of Facebook usage in an ethnically polarized setting. *Proceedings of the National Academy of Sciences* **118**, e2022819118. doi:10.1073/pnas.2022819118
- Baker P and Potts A (2013) 'Why do white people have thin lips?' Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies* **10**, 187–204. doi:10.1080/17405904.2012.744320

- Bakir V and McStay A** (2018) Fake news and the economy of emotions. *Digital Journalism* **6**, 154–175. doi:10.1080/21670811.2017.1345645
- Balkin JM** (2008) The constitution in the national surveillance state. *Minnesota Law Review* **93**, 1–25.
- Berger J and Milkman KL** (2012) What makes online content viral? *Journal of Marketing Research* **49**, 192–205.
- Brady WJ, Wills JA, Jost JT, Tucker JA and Van Bavel JJ** (2017) Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* **114**, 7313–7318. doi:10.1073/pnas.1618923114
- Bursztyjn L, Egorov G, Enikolopov R and Petrova M** (2019) *Social media and xenophobia: evidence from Russia*. Tech. Rep. National Bureau of Economic Research.
- Deibert RJ** (2019) Three painful truths about social media. *Journal of Democracy* **30**, 25–39. doi:10.1353/jod.2019.0002
- Diakopoulos N** (2015) Algorithmic accountability. *Digital Journalism* **3**, 398–415. doi:10.1080/21670811.2014.976411
- Dommett K and Power S** (2019) The political economy of Facebook advertising: election spending, regulation and targeting online. *The Political Quarterly*. doi:10.1111/1467-923x.12687
- Eckles D, Gordon BR and Johnson GA** (2018) Field studies of psychologically targeted ads face threats to internal validity. *Proceedings of the National Academy of Sciences* **115**, E5254–E5255. doi:10.1073/pnas.1805363115
- Edelson L and McCoy D** (2021) *We research misinformation on Facebook. It just disabled our accounts*. Available at <https://www.nytimes.com/2021/08/10/opinion/facebook-misinformation.html>, accessed 4 August 2021.
- Eslami M, Rickman A, Vaccaro K, Aleyasen A, Vuong A, Karahalios K, Hamilton K and Sandvig C** (2015) "I always assumed that I wasn't really that close to [her]" reasoning about invisible algorithms in news feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 153–162.
- Faizullahbhoj I and Korolova A** (2018) *Facebook's advertising platform: new attack vectors and the need for interventions*. CoRR, abs/1803.10099. Available at <http://arxiv.org/abs/1803.10099>
- Fisher M, Cox JW and Hermann P** (2016) *Pizzagate: from rumor, to hashtag, to gunfire in DC*. Available at https://www.washingtonpost.com/local/pizzagate-from-rumor-to-hashtag-to-gunfire-in-dc/2016/12/06/4c7def50-bb4d-11e6-94ac-3d324840106c_story.html, accessed 13 April 2020.
- Freedom House**. (2020). *Freedom in the world 2020. A leaderless struggle for democracy*. Tech. Rep.
- Garcia D** (2017) Leaking privacy and shadow profiles in online social networks. *Science Advances* **3**, e1701172. doi:10.1126/sciadv.1701172
- Giroux HA and Bhattacharya D** (2016) Anti-politics and the scourge of authoritarianism. *Social Identities*. doi:10.1080/13504630.2016.1219145
- Heawood J** (2018) Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity* **23**, 429–434. doi:10.3233/IP-180009
- Hills TT** (2019) The dark side of information proliferation. *Perspectives on Psychological Science* **14**, 323–330. doi:10.1177/1745691618803647
- Hills TT, Noguchi T and Gibbert M** (2013) Information overload or search-amplified risk? Set size and order effects on decisions from experience. *Psychonomic Bulletin & Review* **20**, 1023–1031. doi:10.3758/s13423-013-0422-3
- Jost JT, Barberá P, Bonneau R, Langer M, Metzger M, Nagler J, Sterling J and Tucker JA** (2018) How social media facilitates political protest: information, motivation, and social networks. *Political Psychology* **39**, 85–118. doi:10.1111/pops.12478
- Kaiser J and Rauchfleisch A** (2020) Birds of a feather get recommended together: algorithmic homophily in YouTube's channel recommendations in the United States and Germany. *Social Media+Society* **6**. doi:10.1177/2056305120969914
- Keller M** (2013) *The Apple 'kill list': what your iPhone doesn't want you to type*. Available at <https://www.thedaily-beast.com/the-apple-kill-list-what-your-iphone-doesnt-want-you-to-type>, accessed 20 April 2020.
- Kozyreva A, Lewandowsky S and Hertwig R** (2020) Citizens versus the Internet: confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest* **21**, 103–156. doi:10.1177/1529100620946707
- Kozyreva A, Lorenz-Spreen P, Hertwig R, Lewandowsky S and Herzog SM** (2021) Public attitudes towards algorithmic personalization and use of personal data online: evidence from Germany, Great Britain, and the United States. *Humanities and Social Sciences Communications* **8**. doi:10.1057/s41599-021-00787-w
- Lewandowsky S, Smillie L, Garcia D, Hertwig R, Weatherall J, Egidy S, Robertson RE, O'Connor C, Kozyreva A, Lorenz-Spreen P, Blaschke Y and Leiser M** (2020) *Technology and democracy: understanding the influence of online technologies on political behaviour and decision making*. Tech. Rep. doi:10.2760/709177
- Lorenz-Spreen P, Mønsted BM, Hövel P and Lehmann S** (2019) Accelerating dynamics of collective attention. *Nature Communications* **10**, 1759. doi:10.1038/s41467-019-09311-w
- Lorenz-Spreen P, Lewandowsky S, Sunstein CR and Hertwig R** (2020) How behavioural sciences can promote truth and, autonomy and democratic discourse online. *Nature Human Behaviour* **4**, 1102–1109. doi:10.1038/s41562-020-0889-7

- Lührmann A and Lindberg SI** (2020) *Autocratization surges – resistance grows*. Democracy Report 2020 (Tech. Rep.). Gothenburg: V-Dem Institute.
- Matz SC, Kosinski M, Nave G and Stillwell DJ** (2017) Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences* **48**, 12714–12719. doi:10.1073/pnas.1710966114
- Merrill JB and Oremus W** (2021) *Five points for anger, one for a 'like': how Facebook's formula fostered rage and misinformation*. Available at <https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/>, accessed 31 October 2021.
- Müller K and Schwarz C** (2019) Fanning the flames of hate: social media and hate crime. *SSRN Electronic Journal*. doi:10.2139/ssrn.3082972
- Noble SU** (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: New York University Press.
- Pariser E** (2011) *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.
- Paschen J** (2019) Investigating the emotional appeal of fake news using artificial intelligence and human contributions. *Journal of Product & Brand Management* **29**, 223–233. doi:10.1108/jpbm-12-2018-2179
- Pasquale F** (2015) *The Black Box Society*. Cambridge, MA: Harvard University Press.
- Persily N** (2017) Can democracy survive the Internet? *Journal of Democracy* **28**, 63–76.
- Pothos EM, Lewandowsky S, Basieva I, Barque-Duran A, Tapper K and Khrennikov A** (2021) Information overload for (bounded) rational agents. *Proceedings of the Royal Society B: Biological Sciences* **288**, 20202957. doi:10.1098/rspb.2020.2957
- Powers E** (2017) My news feed is filtered? *Digital Journalism* **5**, 1315–1335. doi:10.1080/21670811.2017.1286943
- Rader E and Gray R** (2015) Understanding user beliefs about algorithmic curation in the Facebook news feed. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. doi:10.1145/2702123.2702174
- Ricci F, Rokach L and Shapira B** (2015) *Recommender Systems: Introduction and Challenges*. New York, NY: Springer.
- Rød EG and Weidmann NB** (2015) Empowering activists or autocrats? The Internet in authoritarian regimes. *Journal of Peace Research* **52**(3), 338–351. doi:10.1177/0022343314555782.
- Schaub M and Morisi D** (2020) Voter mobilisation in the echo chamber: broadband Internet and the rise of populism in Europe. *European Journal of Political Research*. doi:10.1111/1475-6765.12373
- Soroka S, Fournier P and Nir L** (2019) Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences* **116**, 18888–18892. doi:10.1073/pnas.1908369116
- Starke C, Naab T and Scherer H** (2016) Free to expose corruption: the impact of media freedom, Internet access and governmental online service delivery on corruption. *International Journal of Communication* **10**, 4702–4722.
- Susser D, Roessler B and Nissenbaum H** (2019) Online manipulation: hidden influences in a digital world. *Georgetown Law Technology Review* **4**, 1–45.
- Sweeney L** (2013) Discrimination in online ad delivery. *Queue* **11**, 1–19. doi:10.1145/2460276.2460278
- Tucker JA, Theocharis Y, Roberts ME and Barberá P** (2017) From liberation to turmoil: social media and democracy. *Journal of Democracy* **28**, 46–59. doi:10.1353/jod.2017.0064
- Van Bavel JJ, Rathje S, Harris E, Robertson C and Sternisko A** (2021) How social media shapes polarization. *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2021.07.013
- Wu T** (2017) *The Attention Merchants*. London, UK: Atlantic Books.
- Youyou W, Kosinski M and Stillwell D** (2015) Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences* **112**, 1036–1040. doi:10.1073/pnas.1418680112
- Zuboff S** (2019) Surveillance capitalism and the challenge of collective action. *New Labor Forum* **28**, 10–29. doi:10.1177/1095796018819461

Stephan Lewandowsky is a cognitive scientist at the University of Bristol. His research focuses on people's responses to misinformation and the potential tension between online technology and democracy.

Peter Pomerantsev is a senior fellow at the SNF Agora Institute at Johns Hopkins University where he co-directs the Arena Initiative, a research project dedicated to overcoming the challenges of digital era disinformation and polarisation.