Democracy@Risk is a research project led by Professor Emma Barrett and Professor Rachel Gibson at the University of Manchester. The project is focused on identifying, measuring and responding to the new threats (and opportunities) that the deployment of digital technologies and AI pose to the functioning of democratic institutions and the deeper cultural norms of trust that underpin them. Democracy@Risk

- DIL-related skills and knowledge are dispersed, disconnected and rarely applied to the modern digital world under the current curriculum. The burden of realising the civic importance of DIL, making DIL-related connections across subjects, understanding how these apply to the digital information landscape, and subsequently practising this, lies almost entirely on students.

- There exist no systemic DIL-related training, standards and support for teachers, meaning that current provisions of DIL in schools are the province of individual teachers who recognise the importance of DIL education independent of curricular standards. As such, DIL provisions are likely to vary both within and between schools,

FIGURE 1
Digital Political Micro-Targeting in numbers

43% of total advertising spending by political campaigns in the UK was spent on digital advertising in 2017 [4]

5.9 million Facebook ads run by the Trump campaign in 2016 [6]

$192.3 million spent by Trump and Biden campaigns on Facebook ads during Jan-Oct 2020 [5]

€17.3 million spent on 185,988 Google political ads in the EU and UK since March 2019 [7]

Common concerns

Common concerns about digital political micro-targeting can be divided into three broad types:

1. Persuasive capabilities:

In the wake of the Cambridge Analytica scandal, concerns have been raised over whether political campaigns are becoming too 'scientific', with sensitive personal data leaving voters vulnerable to 'subconscious' forms of influence and manipulation. In particular, many question whether 'psychographic' targeting could be used to tap into hitherto-unavailable forms of subliminal influence, manipulate voter opinions and subsequently 'steal' elections.

2. Regulatory:

The second set of concerns questions whether digital political micro-targeting meets existing standards for political advertising which were created for traditional forms of communication, and focuses on the extent to which we are able to monitor this practice and check compliance. Many worry that this practice represents a 'wild west' in political advertising that falls short of the established standards for transparency and privacy and facilitates loopholes in electoral regulation.

3. Structural:

As digital political micro-targeting becomes more widespread, some scholars and observers are concerned that this could bring about a hyper-segmentation of the electorate that would transform political communication and public discourse, and incentivise political behaviour that could ultimately harm relationships between candidates and voters.

In the next section, we discuss the extent to which these concerns appear to be supported by findings in extant research and identify the resulting implications for democracy. Whist we find concerns over persuasive capabilities to be the least supported by research, regulatory and structural concerns have a stronger evidential basis, with significant challenges for democracy emerging on these fronts.

Can digital micro-targeting 'steal' elections through subliminal persuasion?

It is unlikely that elections can be 'stolen' through the subtle manipulation of voter opinions via digital political micro-targeting for the following reasons:

1. Micro-targeted political communication carries low persuasive power

Whilst personality has been linked to voter choices[8] and psychological targeting does appear to have some sway over consumer behaviour,[9] we find little empirical support for the claim that 'psychographic' micro-targeting is a powerful method of persuading voters to abandon or alter their pre-existing political beliefs and positions.

FIGURE 3
Gaps in oversight

Regulators

- UK electoral law last reformed in 2001, before the era of digital campaigning and emergence of social media platforms [20]

Ad Monitoring

- No common standards for ad libraries and definitions of 'political' ad

- Age, gender and location are the only audience details shown in Facebook ad library

- Google

- Facebook algorithms have been found to inhibit a political campaign's ability to reach voters with diverse views and skew audiences in ways that are unintended by and unapparent to advertisers [25]

Citizen-level Oversight

- Fewer than 15% of data transfers to third-parties are disclosed in website privacy policies [26]

Consequences of tailored communication

1. Particularistic policy promises
Greater reliance on digital micro-targeting puts political parties and candidates at risk of pursuing policy with particularistic benefits, instead of delivering broader public goods.

Scholars have modelled the behaviour of political actors to show that when candidates are able to target messages to specific groups of voters, the result is greater commitment to projects that benefit small groups, even if the social cost of such projects outweighs the benefits.[39] Digital political micro-targeting therefore risks resulting in inefficient policy platforms which do not represent or respond to the needs of the electorate at large.

2. Overemphasis on wedge issues

It is highly likely that greater reliance on digital political micro-targeting will result in the dominance of divisive issues at election time.

An earlier study of traditional 'offline' campaigning in the USA has shown political actors are more likely to focus on 'wedge' issues (divisive topics such as immigration, crime or sexuality which cause conflict within an otherwise united group) in targeted forms of communication, when compared with messages which are more widely broadcast (such as television advertising).[40] Scholars have also constructed models which show that the ability to target messages is likely to result in more extreme positions being taken on wedge issues by political candidates.[41]

Democratic Harms

Weaker representation and collective agenda
The overwhelming focus on wedge issues is not only divisive but also raises considerable problems

Digital political micro-targeting must be made more transparent to enable accountability and maintain public trust. Common industry standards are needed and the balance of regulation must be moved away from tech company self-regulation towards independent monitoring bodies.

Actions can include:

- Updating UK electoral legislation to establish specific rules and reporting processes for digital campaigning and spending

- Creating industry standards for ad libraries. Records should include full details of the targeting, optimisation and ad placement conditions specified by advertisers, and more granular data on spending, impressions and type of political advertiser [42]

- Regular audits of ad delivery algorithms for skews in distribution and possible misalignment with data protection regulation [43]

- Limiting the number of ads a political campaign can run per week to help over-stretched regulators keep up with monitoring ad content in real-time

The ability to distribute more selective and tailored messages encourages political actors to engage in actions that are legal but nevertheless damaging for democracy in the long-term. Improvements in oversight alone will not be able to prevent this problematic political behaviour.

We recommend focusing on deterrence measures which reduce the appeal and raise the costs of problematic political behaviour.

Actions can include:

- Introducing hybridity to digital micro-targeting. For example, introducing broadcasting to micro-targeting by allowing the targeting only of those posts which are published (and permanently archived) on the official social media pages of political parties or candidates.

- Forcing campaigns to issue notices or apologies to those users who have been exposed to ads that breached regulatory standards

- Raising the maximum financial penalties that can be issued by regulators [44]

Empowering citizens can make oversight more multidimensional and build critical resilience that reduces the harmful effects of digital micro-targeting in real-time, particularly if problematic political ads manage to escape the regulatory filter.

Actions can include:

- Enabling users to opt-out of political advertising online

- Providing users with easily-accessible and non-binary control tools over the kind of data that can be used to track and target them

- Investing into digital information literacy programmes for citizens

Who creates online misinformation?

State actors

Intelligence services and state-sponsored agencies in Russia,[48] Iran[49] and China[50] have been found to engage in extensive digital information operations on social media with the aim of influencing public discourse and politics in other states. State actors have also been found to conduct disinformation operations targeted at domestic politics, as with the South Korean National Intelligence Service during the nation's 2012 presidential elections.[51]

Media

Media outlets such as the Russian state-sponsored RT and Sputnik are known to publish information that has been twisted into misleading 'strategic narratives'.[52] Research has also uncovered the publication of inaccurate and misleading information (and the subsequent concealment of this) by the mainstream media in the West, particularly during the early stages of rapidly-developing crises and shock events.[53]

Third-party agents

Media reports show that extensive disinformation operations have originated from teenagers in Macedonia,[54] individuals in Romania,[55] and commercial actors in the USA.[56] US conspiracy theorists have also been identi

Apathy, resignation and withdrawal
Widespread misinformation - or the perception of
such, as a result of cynicism - increases the
cognitive demands for citizens, which can feel
overwhelming and breed paralysing uncertainty.

Apathy, resignation and withdrawal
Widespread misinformation - or the perception of
such, as a result of cynicism - increases the

Can misinformation be corrected?

Whilst fact-checking practices are valuable, they are not enough to protect voters from the influence of misinformation.

On the one hand, corrections have been found to reduce misperceptions and the persuasive power of fake news stories, thus resulting in more accurate beliefs.[76] Yet important caveats remain:

1. Fact-checking organisations have limited reach and resources
The limited financial and labour resources of fact-checking organisations in the face of a vast eco-system of misinformation mean that fact-checkers have to prioritise the investigation of some claims

FIGURE 8
Fact-checking as a filter of misinformation
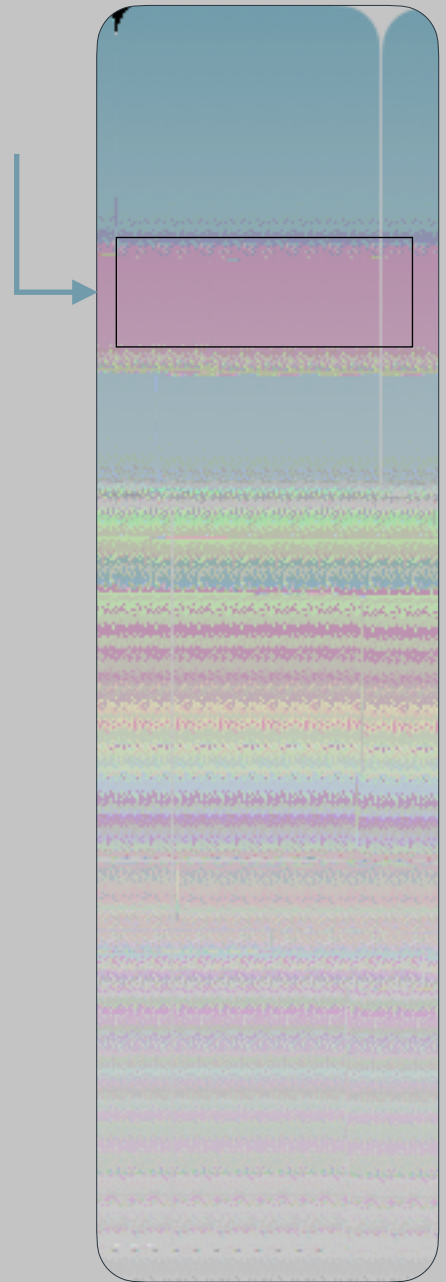
The recommendations outlined in the previous parts of this report touch on the common theme of digital information literacy (DIL). In the fight against online misinformation, DIL represents a   measure that is likely to be rkdayions outliW) 024 (ithmon)] TspecT Q(DIL). poEMic

DIL education in the UK

Current DIL provisions in the UK are facing the following challenges:

1. Absence of an overarching strategic framework for the provision of DIL through primary, secondary and further education

The current curriculum therefore places the burden of connecting subject-specific knowledge and skills, translating them into DIL capabilities and learning to apply these to the digital information landscape entirely on the students.

Figure 10 visualises the above disconnect. To improve provisions for DIL, the national curriculum does not necessarily need to prioritise the introduction of DIL as a separate and new subject - rather, such literacy requires further assembly through the practice of online reasoning in and the fostering of collaboration between existing school subjects.

3. DIL provisions are not underpinned by large-scale systematic support for teachers

Recent findings show that teachers are important gatekeepers in educational innovation, with one study finding that computing teachers are able to reject innovations in digital technology curricula and hinder educational reform in practice, even when this carries a legal mandate and support from industry.[99] As a result, teacher-level factors, including teacher perceptions of the importance of DIL in education and the extent to which they are prepared to incorporate it into their teaching and learning, must be taken into consideration in discussions of DIL education in the UK. Given the multidisciplinary demands of DIL, such considerations should not be limited to computing teachers but instead account for teachers from a wide variety of subject backgrounds.

In the absence of a national DIL framework and standard for assessment, however, we know relatively little about how well teachers are equipped for the delivery of DIL-related skills and knowledge or prepared for educational reform on this front. The partial insights we do possess, however, give cause for concern.

For example, a recent study pointed to surprisingly low levels of DIL among university students and even historians with a PhD, despite their extensive experience as internet users, suggesting that academic excellence and expertise within a subject is not a guarantee of adequate levels of DIL-related skills and knowledge.[100] Teachers should not, therefore, be assumed to possess expertise in DIL

by virtue of their high levels of education and subject specialism.

Teacher training is unlikely to have filled such gaps in the DIL levels of teachers, since current teacher-training programmes (such as the PGCE) do not set out specific requirements for or offer training in DIL as a mandatory component of teacher education.

Whilst numerous organisations have created informative DIL-related teaching resources, these cannot act as substitutes for systematic support and training for two reasons. First, the vast majority of such resources and training opportunities focus on the delivery of one-off (or a short series of) sessions focusing on specific issues such as 'fake news'. As such, they offer little practical pedagogical guidance on how DIL can be embedded into teaching and learning in the long-term and do not help teachers to build the long-lasting multidisciplinary linkages within the curriculum necessary for the effective consolidation of DIL-related skills and knowledge over time. Second, the use of such resources and training opportunities by teachers depends largely on self-referral and therefore depends on teachers' pre-existing recognition of the importance of DIL in education or awareness of the gaps in one's own DIL-related competencies. Thus, there is no guarantee that the majority of teachers will actively seek out and translate these opportunities into teaching practice.

In the absence of coherent DIL standards and mandatory training for teachers, the capacity and willingness of individual teachers to foster among their students the full set of skills, knowledge and attitudes encompassed by DIL is likely to be inconsistent at best - or very low at worst. Without the introduction of large-scale systematic DIL-related incentives and support for teachers, unequal provisions of DIL within and between schools emerge as the best case scenario for schools in the UK.

D igital technology is making the world better connected - it has never been cheaper, quicker or easier to locate, create and share information on a mass scale. As a result, the digital age is seeing the overwhelming majority people conduct key aspects of their lives online and for longer periods of time. This increased connectivity and digital presence, however, mean that access to the very core of power in democracies - the citizens - is widening and, if left unchecked, risks exposing democracy to multiple intended and unintended harms.

The multidisciplinary review conducted by the Democracy@Risk Project sought to offer more clarity with respect to two areas of such harm - digital political micro-targeting and online misinformation. As summarised in this report, the extant research and scholarship suggest that these forms of activity generate considerable risks for democracy - consequences that are often, albeit not always, unforeseen by those involved.

Our review found that the biggest sources of risks extend beyond the   power of the information that enters citizen consciousness through digital micro-targeting or the eco-system of online misinformation. Elections have been characterised by imperfect information and cognitive biases have clouded the political judgements of citizens long before the digital era, and the digital age has not generated the technology that can pierce through these challenges and mould voter opinions with certainty.

Other processes and principles fundamental to the effective functioning of democracy, however, are more at risk. The first of these is
                  , which requires for all citizens to have equal rights to participate in democratic discourse and decision-making, as well as equal protections of those rights. Digital political micro-targeting, for example, hinders the ability of our democratic political system to bring everyone to the table both unintentionally, by enabling political redlining, and intentionally, by facilitating voter suppression efforts. Online misinformation can push citizens further away by generating disempowering uncertainty that encourages resignation and avoidance of politics.

Second,
processes are at risk. Greater reliance on digital

micro-targeting is likely to result in more fragmented

During a special launch event in September 2021, we received valuable feedback and further comments on the themes covered in this report from three leading experts[101] in the fields of political micro-targeting, online misinformation and digital information literacy.

This final section seeks to share with the wider research community the key insights we took away from this event, in the hope that this might help to identify some further priorities for policy-oriented research that build on the work conducted by the Democracy@Risk project thus far.

Recommendations:

1. Assessments of potential harms to democracy should apply democratic theory to identify a workable, unifying vision of democracy as a political system. Future research should help policy-makers to relate specific forms of democratic harm (such as 'reduced

11.

27. Costa, E. & Halpern, D. (2019). The behavioural science of online harm and manipulation, and what to do about it, March 2019 [Online] Available from https://www.bi.team/wp-content/uploads/2019/04/BIT_The-behavioural-science-of-online-harm-and-manipulation-and-what-to-do-about-it_Single.pdf [Accessed April 2020]

28. Libert, T. (2018). An automated approach to auditing disclosure of third-party data collection in website privacy policies.   April 23-27, 2018. Lyon, France.
as cited in Costa, E. & Halpern, D. (2019). The behavioural science of online harm and manipulation, and what to do about it,   March 2019 [Online] Available from https://www.bi.team/wp-content/uploads/2019/04/BIT_The-behavioural-science-of-online-harm-and-manipulation-and-what-to-do-about-it_Single.pdf [Accessed April 2020]

29. Open Rights Group. (2020). Public are kept in the dark over data driven political campaigning, poll finds.   10 January. Available from https://www.openrightsgroup.org/press-releases/public-are-kept-in-the-dark-over-data-driven-political-campaigning-poll-finds/ [Accessed June 2021]

30. Mozilla. (2019). Data Collection Log — EU Ad Transparency Report,   Available from https://adtransparency.mozilla.org/eu/log/ [Accessed May 2020]
Mozilla. (2019). Facebook and Google: This is What an Effective Ad Archive API Looks Like,   Available from https://blog.mozilla.org/en/mozilla/facebook-and-google-this-is-what-an-effective-ad-archive-api-looks-like/ [Accessed June 2021]
Who Targets Me. (2020).How to take a "gold standard" approach to political advertising transparency and policy,   Available from https://whotargets.me/en/how-to-take-a-gold-standard-approach-to-political-advertising-transparency-and-policy/ [Accessed June 2021]

31. Facebook. (2019). Boost liquidity and work smarter with machine learning,   27 March. [Online] Available from https://www.facebook.com/business/news/insights/boost-liquidity-and-work-smarter-with-machine-learning [Accessed June 2020].

42. Examples of more detailed recommendations for ad library standards can be found in:
Who Targets Me. (2020).How to take a "gold standard" approach to political advertising transparency and policy, Available from https://whotargets.me/en/how-to-take-a-gold-standard-approach-to-political-advertising-transparency-and-policy/
Mozilla. (2019). Facebook and Google: This is What an Effective Ad Archive API Looks Like, Available from https://blog.mozilla.org/en/mozilla/facebook-and-google-this-is-what-an-effective-ad-archive-api-looks-like/ [Accessed June 2021]

43. More detailed recommendations for auditing tools can be found in:
Who Targets Me. (2020).How to take a "gold standard" approach to political advertising transparency and policy, Available from https://whotargets.me/en/how-to-take-a-gold-standard-approach-to-political-advertising-transparency-and-policy/ [Accessed June 2021]

44. More detailed recommendations on increasing financial sanctions can be found in:
APPG on Electoral Campaigning Transparency. (2020). Defending our democracy in the digital age. [Online]. Available from https://fairvote.uk/defending-our-democracy-in-the-digital-age-new-report-launched/ [Accessed April 2020].
The Electoral Commission. (2021). Reforming electoral law. [Online]. Available from https://www.electoralcommission.org.uk/sites/default/files/pdf_file/Reforming-electoral-law-PACAC-booklet.pdf [Accessed June 2021]

45. Innes, M. (2020). Techniques of disinformation: Constructing and communicating "soft facts" after terrorism, (2). 284-299.

46. Nielsen, R.K., Fletcher, R., Newman, N., Brennen, J.S., & Howard, P. (2020). Navigating the 'Infodemic': How People in Six Countries Access and Rate News and Information about Coronavirus. Reuters Institute for the Study of Journalism, April 2020. Available from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-04/Navigating%20the%20Coronavirus%20Infodemic%20FINAL.pdf [Accessed May 2020]

47. Innes, M. (2020). Techniques of disinformation: Constructing and communicating "soft facts" after terrorism, (2). 284-299.

48. Howard et al. (2018). The IRA and Political Polarization in the United States, , August 22, 2018. Available from: https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/12/The-IRA-Social-Media-and-Political-Polarization.pdf [Accessed April 2020].
DiResta, R., et al. (2018). The Tactics & Tropes of the Internet Research Agency, , December 17, 2018. Available from https://cdn2.hubspot.net/hubfs/4326998/ira-report-rebrand_FinalJ14.pdf [Accessed 10 April 2020]

49. Nimmo, B., Eib, S., Ronzaud, L., Ferreira, R., Lederer, T., & Smith, M. (2020). Iran's Broadcaster: Inauthentic Behavior , 5th May 2020. Available from https://www.graphika.com/reports/irans-broadcaster-inauthentic-behavior/ [Accessed May 2020]

50. Nimmo, B., Francois, C., Eib, S. & Ronzaud, L. (2020). Return of the (Spamouflage) Dragon, , 24th April 2020. Available from https://www.graphika.com/reports/return-of-the-spamou

55.

71. Huang, H. (2017). A War of (Mis)Information: The political effects of rumors and rumor rebuttals in an authoritarian country,

84. McGuire, W. J. (1961). The effectiveness of supportive and refutational defenses in immunizing and restoring beliefs against persuasion, , 184–197.
McGuire, W. J. & Papageorgis, D. (1961). The relative efficacy of various types of prior belief-defense in producing immunity against persuasion, J

House of Lords Select Committee on Democracy and Digital Technologies. (2020).

(CP 285). [Online]. London: House of Lords. Available from https://committees.parliament.uk/publications/2308/