

Toward an integrated and globally equitable understanding of political polarization

Received: 25 July 2025

Accepted: 16 July 2026

Published online: 02 September 2026

 Check for updates

Max Falkenberg^{1,2}, Matteo Cinelli³, Alessandro Galeazzi⁴, Christopher A. Bail⁵, Rosa M. Benito⁶, Axel Bruns⁷, Anatoliy Gruz⁸, David Lazer^{9,10}, Jae Kook Lee¹¹, Jennifer McCoy¹², Kikuko Nagayoshi¹³, David G. Rand¹⁴, Antonio Scala³, Alexandra A. Siegel¹⁵, Sander van der Linden¹⁶, Onur Varol¹⁷, Ingmar Weber¹⁸, Magdalena Wojcieszak^{19,20}, Fabiana Zollo²¹, Andrea Baronchelli²²✉ & Walter Quattrociocchi³✉

Political polarization poses major societal risks, yet a globally relevant, interdisciplinary understanding remains lacking, as nearly half of all studies focus on the United States. We argue for renewed effort to bring global equity to polarization research, highlighting insights robust across countries, those unique to specific contexts, and key literature gaps. Closing these gaps means overcoming cultural and systemic barriers, including data-access restrictions and misaligned incentives. It also demands interdisciplinarity bridging traditional approaches and those addressing the polarizing role of the internet, social media, and AI. Otherwise, efforts to counter polarization, and its democratic harms, risk resting on unsuitable evidence.

With a folk understanding that political polarization refers to socio-political divisions within a society, many have proclaimed that we are more divided than ever^{1–3}. In this account, polarization has been blamed for populism⁴, the erosion of social cohesion and rise of incivility^{5–9}, the loss of trust in the institutions of democracy^{10,11}, legislative dysfunction^{11–14}, and the collective failure to address existential risks such as Covid-19 or climate change^{15–17}.

But political polarization is far from a new problem—it has been a key focus of research since the birth of the social sciences—and it is reasonable to argue that, when viewed in a historical context, current societal divisions are minor compared to those which led to the major global conflicts of the 20th century¹⁸. What has fundamentally changed is the emergence of the information age (the internet, social media, AI) and its integration into, and domination over, our contemporary

political discourse¹⁹. This shift, and the election of Donald Trump to the U.S. Presidency in 2016 on the back of a campaign that heavily leveraged social media, has driven a renewed interest in understanding the causes and consequences of political polarization. However, while polarization research has blossomed in the U.S., a comparable evidence-base for understanding polarization in a global context remains limited (Fig. 1; see Supplementary Note 1 for data collection details), especially outside Europe and other high-income democracies. If insights from the U.S. were to generalize to other regions, the lack of localized evidence could, perhaps, be tolerated. However, as we will outline, there are numerous results from the U.S. which are outliers when viewed in a global context, where there are substantial conceptual challenges which threaten our normative understanding of what polarization is and how it should be addressed (Box 1; see ref. 20).

¹King's College London, London, UK. ²Central European University, Vienna, Austria. ³Sapienza University of Rome, Rome, Italy. ⁴University of Padova, Padova, Italy. ⁵Duke University, Durham, NC, USA. ⁶Universidad Politécnica de Madrid, Madrid, Spain. ⁷Queensland University of Technology, Brisbane, Australia. ⁸Toronto Metropolitan University, Toronto, Canada. ⁹Northeastern University, Boston, MA, USA. ¹⁰Harvard University, Cambridge, MA, USA. ¹¹Sungkyunkwan University, Seoul, South Korea. ¹²Georgia State University, Atlanta, GA, USA. ¹³University of Tokyo, Tokyo, Japan. ¹⁴Cornell University, Ithaca, NY, USA. ¹⁵University of Colorado Boulder, Boulder, CO, USA. ¹⁶University of Cambridge, Cambridge, UK. ¹⁷Sabancı University, İstanbul, Türkiye. ¹⁸Saarland University, Saarbrücken, Germany. ¹⁹University of Warsaw, Warsaw, Poland. ²⁰University of California Davis, Davis, CA, USA. ²¹Ca' Foscari University of Venice, Venice, Italy. ²²City St. George's University of London, London, UK. ✉e-mail: andrea.baronchelli.1@citystgeorges.ac.uk; walter.quattrociocchi@uniroma1.it

BOX 1

The normative U.S. framing of polarization

In selecting “polarization” as their word of 2024, Merriam-Webster defined polarization as “division into two sharply distinct opposites; especially: a state in which the opinions, beliefs, or interests of a group or society no longer range along a continuum but become concentrated at opposing extremes”¹⁹⁷. While such a framing may be appropriate for the public’s understanding of the U.S.’s unusual Manichean two-party democracy, there are many cases where such a framing may not be appropriate²⁰. For instance, it is far from clear how polarization—conceptualised in a variety of ways (ideological, affective), and differently by different research fields (e.g., distributions of opinions in surveys, politically homophilous interactions on social media⁹)—should be conceptualised²⁰, and measured, in multi-party democracies.

Researchers have adapted conventional two-party measures of polarization for use in multi-party contexts^{7,20,70,74,78}, but, to some extent, it remains an open question as to whether key concepts in the study of polarization change when extended from a one-dimensional political spectrum to a multi-dimensional context. Key areas of dispute include the importance of in-group identification in defining affective polarization⁷⁸, the distinction between animosity aimed at political elites⁵⁶, partisans, and voters more generally, the precise nature of political animosity²⁰, and whether we should expect multiple parties to fall into two, or more, opposed camps^{20,78}.

Similarly, there is substantial work discussing how polarization can be formulated without explicit reference to the extremes, e.g., by referencing the bimodality or the spread of the left-right distribution, or without reference to political ideology, changing how the concept is interpreted⁴⁴. While such issues can seem abstract, they can directly affect the global context for understanding polarization when, for example, cross-country political values surveys ask individuals to place themselves ideologically on a one-dimensional left-right scale, but not on other axes of political belief which may be equally important for understanding political polarization (e.g., the establishment/anti-establishment dimension^{15,198,199}).

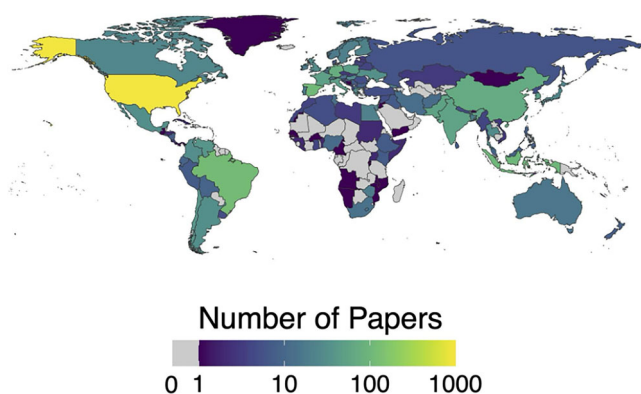


Fig. 1 | The unequal distribution of polarization research. Academic papers on political polarization which refer to a specific country overwhelmingly reference the United States. In total, 5965 papers use the term “political polarization”, “ideological polarization” or “affective polarization” (or equivalent spelling in British English) in their title or abstract of which 2243 refer to at least 1 country. The United States is mentioned in 967 papers (43.1% of papers with a country mention). Many papers which do not mention a country are U.S. focused, but do not state so explicitly. Contextual factors which contribute to the unequal distribution of polarization research include, but are not limited to, legitimate global interest in the polarization of U.S. politics given its disproportionate influence on global affairs, differences in academic research funding across countries, the challenges of conducting studies in countries with “hard to reach populations”, and the perception that studying polarization in countries where polarization is particularly high, or damaging, is more important than in less polarized countries.

Ensuring a globally relevant literature on political polarization is, or should be, a fundamental goal for the academic community (as several have argued previously^{1,7,8,21,22}). But there are substantial systemic and cultural barriers to achieving this goal. First, how polarization is studied is often influenced by norms established in U.S.-focused research (Box 1). Second, prominent reviews^{6,23–25}, and research articles^{26–29}, published in top interdisciplinary journals are often limited to the U.S., which may in turn shape academic incentives. Finally, when research does focus on non-U.S. contexts, the approach employed is usually confined to limited disciplinary silos, either employing traditional social science research methods (often using survey data), or applying data- and computational social science

approaches to study polarization in our contemporary information ecosystem (e.g., on social media). Rarely are these streams of research synthesized into a single coherent narrative and understanding which is able to capture the complex societal feedback loops across the different domains (political institutions, the economy, legacy and social media) which contribute to driving and sustaining polarization. Quantitative polarization research is contingent on high-quality data. But in the case of digital media research, access to this data has become increasingly difficult (an issue which the European Union’s Digital Services Act aims to fix^{30–32}), and is sometimes only available to U.S.-based researchers (for example, in the case of Meta’s U.S. election studies^{27,28,33–35}).

Seeking to address these issues, this article offers a broad overview of polarization research globally, focusing on key trends in ideological and affective polarization (the most widely studied polarization sub-types), the influence of socio-economic factors, the impact of the traditional news media and social media, and the increasingly important role of recommendation systems and artificial intelligence (AI)^{36–38}. Drawing from a growing set of examples, we discuss key results in the literature that are country-specific and those that, broadly speaking, generalize across a wide range of nations. We also highlight key areas where globally relevant literature is lacking and provide concrete suggestions for advancing towards a more globally equitable understanding of polarization.

Changes to information production, dissemination, and consumption have changed political discourse at a global scale, and an understanding and analysis of these factors, and their global variability, must be integrated into any coherent understanding and analysis of the causes and consequences of polarization in the future. A failure to address these gaps is not only potentially dangerous, risking inefficient or counterproductive polarization interventions, but is intellectually complacent, further deepening existing inequities in our global understanding of the phenomenon.

Our review focuses on cross-national patterns in ideological and affective polarization, on the role of socio-economic and institutional factors, and on the impact of the news media, social media, and AI; many adjacent topics necessarily fall outside its scope. Parts of the literature we discuss will seem disconnected, and some may argue that it would be conceptually clearer to present these streams of research in separate articles. However, it is our view that the tendency to study aspects of polarization in a siloed manner only perpetuates this

BOX 2**Challenges in conceptualising and studying polarization in authoritarian states**

Despite its normative framing as a democratic phenomenon, polarization is an important concept for authoritarian states^{52,53,200–202}. As in democracies, individuals in authoritarian societies have diverse ideological views and may have positive or negative affect towards various groups (e.g., towards supporters of the regime in power). But studying polarization in these contexts involves unique challenges (see below). Beyond its inherent value, the study of polarization in these states can offer valuable insights for new democracies which retain a legacy of authoritarianism (e.g., in Kenya²⁰³ and Liberia²⁰⁴) and can guide scholars interested in polarization's role in eroding democracy^{5,60}.

Particularly in authoritarian states, polarization can be exploited as a tool to cement power and discredit opponents²⁰⁵. Researchers can play an important role in understanding political dynamics in these contexts. For example, research has shown how the increased polarization of online interaction patterns preceded periods of violence during the Arab Spring¹²⁹, how the Chinese government produced up to half a billion social media posts each year to distract citizens from political topics²⁰⁶, and how politicians in Zimbabwe deliberately exploited polarization to reshape boundaries of conflict and mobilize constituencies²⁰¹.

Crucially, given the risks posed to research subjects in authoritarian states—including severe physical repression²⁰², or exile²⁰⁷—researchers must address a number of practical and ethical challenges to studying polarization:

- Respondents in authoritarian contexts are “hard to survey”²⁰⁸, meaning that studies suffer from numerous biases, including the censorship of data (e.g., in China¹⁷⁹), preference falsification in surveys, and coordinated inauthentic activity. These biases may be mitigated by using list experiments or conjoint experiments²⁰⁹. Additionally, preference falsification may be counteracted using “anonymity sampling”²¹⁰ which improves survey accuracy while protecting respondent anonymity.
- Scientists are increasingly emphasizing the need for publicly available data for study replication. However, in authoritarian contexts, such data may put survey respondents at risk if not carefully anonymised. Researchers must take extra care when providing public replication datasets—including the IDs of social media posts—that contain politically sensitive content. The use of secure data archives may offer a path for data sharing without putting research subjects at risk²¹¹.
- Even if data are anonymised, scholars should be conscious that their work may be used as “opposition research”, and indirectly benefit the regime in power. For example, papers with real-time monitoring of opposition groups could be used as justification for cracking down on protesters. An example which highlighted these issues is the ENCORE study, in which internet censorship was measured at the user level across a number of countries—including China, Iran, Pakistan, and Saudi Arabia—but without individual consent²¹².

disconnect. For this reason, our review retains a broad scope in the topics covered, pointing to additional resources where relevant, with the overarching aim of fostering greater interdisciplinary cohesion in future polarization research.

Global trends in polarization

Until the early 21st century, scholars across many world regions focused largely on the study of ideological polarization, typically understood as the gap between liberals and conservatives, and by extension between the political left and right^{39–43}. Most commonly, ideological polarization has been quantified at the level of political elites or the mass public according to the dispersion of perceived (in the case of elites) or self-reported (mass public) left-right scores using surveys⁴⁴. Although not aligned a priori, these ideological labels are often considered indicative of specific positions or attitudes on key political issues (e.g., in relation to the economy, immigration and social issues). However, since the 2010s, evidence from the U.S. has called into question this association, noting that the animosity between partisan groups has grown disproportionately relative to policy differences. This has led researchers to shift focus away from ideological polarization and towards affective polarization^{6,23,45,46}, the phenomenon whereby an individual has negative feelings towards their political out-group (and/or positive feelings towards their in-group). Although approaches vary, affective polarization is most commonly measured by employing “feeling thermometers” which gauge partisans’ warmth (or animosity) towards their political opponents⁶.

Ideological polarization

Amongst political elites, comparative data from 21 high-income countries reveal a small but significant increase in ideological polarization since the 1990s. In absolute terms, Canada and Australia are cited as countries with low ideological polarization, in contrast to

Spain, Italy and Sweden where ideological polarization is particularly high⁴⁷. Similarly, across 20 democracies in Western Europe, research finds that the dispersion of political parties along the left-right ideological axis has increased between 1945 and 2021, with researchers suggesting that this has contributed to a fragmented and volatile political context, at the risk of democracy⁴⁸. But such trends in the polarization of political elites do not necessarily reflect growing ideological polarization in the general population.

At a global scale, ideological polarization has not increased across the mass public⁴⁹, which is in clear contrast to the trend observed in the United States, where growing ideological polarization (both mass and elite) is now well documented^{13,50}, albeit this increase is smaller than the variability observed across countries^{6,45,46}.

However, the appropriateness of the left-right ideological scale for assessing political beliefs, and therefore the appropriateness of the concept for measuring polarization, has been disputed in low- and middle-income countries. Here, research suggests that ideological labels do not translate coherently from country to country⁵¹. This lack of coherence also applies in authoritarian contexts, which pose a unique challenge to studying polarization (Box 2). For example, researchers have suggested that the traditional left-right divide, commonly observed in Western political systems, is reversed in China⁵². Here, the lack of a pluralist party system means that ideological divides cannot be mapped onto partisan divides. Researchers have also shown that political attitudes do not map consistently to pro- and anti-regime groups⁵³. These ambiguities challenge how to understand ideological polarization in different national contexts, and especially in authoritarian states where the concept remains understudied⁵³.

Affective polarization

In the United States, the state of affective polarization has been described as “political sectarianism”²³, drawing parallels to religious

and ethnic sectarianism, and having often been linked to the rise of populism (although there is substantial variability in the ideological positions of populist parties globally, see ref. 54). Referencing data from 12 OECD countries, Finkel et al. argue that the state of U.S. affective polarization is exceptional, with the U.S. having suffered the largest increase over the last 40 years⁵⁵. However, estimates of the absolute levels of affective polarization from 53 countries point towards Türkiye, Hungary, and Kenya as examples of high affective polarization, and the Philippines, the Netherlands, Belgium, and Japan as examples of countries with low levels of affective polarization, with the U.S. falling in the middle of the pack⁵ (although this varies depending on whether the focus is on the out-party in general, or on political elites; animosity towards political elites is disproportionately high in the U.S., see ref. 56).

For many countries (including authoritarian states where measuring polarization is particularly challenging) quantitative estimates of affective polarization are still unavailable⁵. In these contexts, where populations are hard to survey (see Box 2), researchers have proposed using expert estimates of polarization, and other markers of democracy, as an alternative to conventional survey studies. The Varieties of Democracy (V-Dem) collaboration⁵⁷ is probably the best-known example of such an international effort. However, it is important to note that expert estimates remain subjective and are open to criticism. In particular, scholars have noted that V-Dem indicators often fail to capture evolving trends, with changes in V-Dem indicators often lagging behind those derived from social surveys⁵⁸. Additionally, the subjectivity of evaluations may result in confirmation bias amongst expert annotators⁵⁹.

Using V-Dem data, researchers have defined the concept of pernicious polarization as quantifying the extent to which a country is divided into mutually antagonistic camps⁶⁰. This measure points towards Brazil, Mexico, the United States, and India as examples of democracies with particularly high levels of pernicious polarization, and reveals that every world region other than Oceania has suffered an increase in pernicious polarization since 2005, albeit levels were much higher in Europe and the Americas during the first half of the twentieth century. However, pernicious polarization is only one of many proposed polarization measures. It is not always clear how different types, or definitions, of polarization are related, in which contexts it is better to study one polarization type over another, and whether polarization, especially ideological polarization, is even undesirable⁶¹.

How are ideological and affective polarization related?

In the United States, the relationship between ideological and affective polarization is debated, rooted in the question of whether policy differences drive partisan animosity, the key component of affective polarization. Some scholars have argued that ideological and affective polarization are largely distinct^{45,46}, while others have identified ideological polarization as a key driver of affective polarization and partisan animosity^{62,63}.

At the heart of this debate is the distinction between two forms of sorting into partisan groups: ideological sorting—the growing alignment of specific issue-positions, and issue-extremity, to partisan affiliation (e.g., Republican or Democrat)⁶²—and social sorting—the alignment between individuals' partisan affiliation and various social identities including race, ethnicity, religion, class, socioeconomic status, location^{46,64,65}, and ideological identity (e.g., liberal or conservative)⁶. In the former case, scholars argue that affective polarization is driven by opposed politicians adopting increasingly divergent issue-positions^{62,66}, with the pairwise alignment of different issues becoming stronger over time (or, alternatively, the diversity of aligned issues broadening; see the “oil-spill” model of polarization⁶⁷). In the latter case, researchers argue that social sorting has resulted in fewer social identities that cut across party lines. Over time, this has resulted in partisans seeing themselves as having less and less in

common with their political opponents, which in turn has strengthened the impact of stereotypes and “othering”, driving affective polarization⁶⁶. In both cases, researchers have argued that the news media and social media play a key role in facilitating growing issue or identity alignment⁶⁸. Here, the crucial and often overlooked observation is that individuals can label themselves as belonging to a specific ideological or partisan group, even if their issue-specific beliefs and attitudes do not appear consistent with those labels (“ideological incongruence”; see ref. 69).

The framing of partisanship in terms of social identity is a crucial change in the conceptual framing of polarization—a feature of politics also observed outside the U.S., including in the UK, Sweden, and the Netherlands⁷⁰. The concept builds on an existing literature from social identity theory which discusses the relationship between group identity and out-group animosity⁷¹, and the role that social identity plays in shaping attitudes during times of political flux (e.g., in Hong Kong in the 1990s⁷²). But the challenge here is that the growing importance of social identity is not easy to disentangle from growing ideological sorting (in the U.S., both have risen in the last 40 years). The consequence is that scholars in the U.S. continue to debate the relationship between ideological and affective polarization due to the lack of a clear causal relationship between the two⁶⁶.

This debate draws parallels globally, where the evidence for the relationship between ideology and affect displays “relatively weak and inconsistent” trends⁷³. Survey evidence suggests that ideological and affective polarization are broadly correlated⁷⁴, but some countries display significant ideological polarization without high out-party animosity (e.g., the Netherlands), while others are afflicted with severe affective polarization despite the principal parties being largely moderate (e.g., Montenegro and Serbia). Importantly, survey evidence from 53 countries suggests that high levels of affective polarization are correlated with democratic backsliding, but high levels of ideological polarization are not⁵ (some previously debated whether ideological polarization is even democratically problematic, since, they argued, some ideological diversity is essential for preventing radicalism⁷⁵). While this comparative evidence is not conclusive, specific examples such as the rise of affective polarization in the UK around Brexit, and associated issues such as immigration, have led scholars to introduce the cross-cutting concept of “issue-based affective polarization”⁷⁶. Scholars have also noted that, in contrast to the conventional framing of affective polarization between partisan groups (e.g., Republicans vs. Democrats), that affective hostility is common across factions within a political party⁷⁷ (e.g., between the progressive and centrist wings of the Labour party in the UK—Corbyn vs. Starmer—or the Democrats in the USA—Sanders vs. Biden). In some instances, the measured level of within-party affective polarization may exceed inter-party polarization⁷⁷, challenging the primacy of partisanship as a political identity.

The reasons for country-level differences in the relationship between ideological and affective polarization are complex and disputed (see²⁰): The U.S.-centric literature has made a two-party “us-vs-them” framing of polarization the norm (Box 1), but in multi-party states it is not uncommon for individuals to have positive (or negative) feelings towards multiple parties (e.g., in Finland or Sweden⁷). This seriously complicates how to measure polarization. For instance, even when affective polarization is measured across multiple parties⁷, the measurement continues to frame the concept in terms of politics “divided into two distinct camps”, a choice criticised by some scholars²⁰. Brazil is a good example of a country which is not well-suited to this formulation of polarization, with politics being less institutionalised and partisan identities being less evident⁷⁸, resulting in affective polarization with low in-group identification. Additionally, framing polarization in a multi-party context allows for new political dynamics which can have both positive and negative democratic impacts, for instance, related to the sustainability and effectiveness of

BOX 3**Do ethnic and religious differences lie at the root of polarization?**

It is often assumed that polarization has roots in ethnic or religious differences. However, while these factors may correlate with observed differences in polarization⁴⁹, multiple comparative studies have shown that social cleavages do not, inherently, predict damaging forms of polarization^{21,60,88,213}.

Ethnic, religious or social class divisions become politically relevant cleavages when people develop shared identities and values, and organizations like parties or unions represent them^{214,215}. And these cleavages become polarizing when political actors make them the basis of polarizing strategies, for instance, through the mobilization of anti-immigration sentiment in Europe and the U.S.²¹⁶. But as Benson²¹⁷ points out, while the social psychology of intergroup conflict may explain the intensification of the in-group identification and out-group animosity of affective polarization, it does not explain how those identities become politically salient in the first place. For this, scholars turn to elite agency and the divisive rhetoric of polarizing politics that politicizes certain social cleavages, grievances and political identities^{201,218}.

For example, ethnic and religious divides were politicized in India and Türkiye as the secular solution of the founders gave way to Hindu nationalism under Modi's BJP in India and the Islamist movement and Kurdish rift under Erdogan's AKP in Türkiye^{219,220}. On the other hand, some countries with significant ethnic divides have not descended into an ethnicity-defined political rift²²¹. A comparative study of four Asian countries finds that those with clearly identifiable social cleavages, Taiwan and Indonesia, suffered only short episodes of damaging polarization, while those with less identifiable social cleavages, Thailand and the Philippines, suffered more destructive consequences, including democratic collapse in the Thai case²²².

Finally, certain cleavages may be more prone to damaging polarization, such as the unresolved, emotionally-charged historic debates around citizenship, national identities, and founding myths that Somer and McCoy²²³ call "formative rifts". Examples include the claim that political legitimacy is to be conferred only on those with national liberation war experience in Zimbabwe²²⁴, or the legacy of unequal citizenship rights and political exclusion around race, gender and religion in the U.S.^{225,226}.

coalition governments⁷⁴. These distinct phenomena are important, since they will contribute to differences in the complex societal feedback loops driving and sustaining polarization.

Differences in the institutional framework of countries are key to understanding cross-national differences in affective polarization and its relationship to ideological polarization. Research suggests that affective polarization is higher in majoritarian political systems⁷³, where political power is concentrated in the hands of a single party (e.g., France, UK) and which feature single-member electoral districts⁷³, than in systems based on consensual governance (e.g., Switzerland, the Netherlands). However, Gidron et al.⁷³ caution against concluding that majoritarian political systems cause affective polarization. In contrast, the ideological polarization of political parties appears lower in more majoritarian electoral systems⁴⁷. As highlighted by ref. 47, one of the best examples of how the political system affects ideological polarization is the example of New Zealand: When the electoral system changed to proportional representation in 1993, the measured level of ideological polarization doubled⁷⁹.

Focusing explicitly on institutions and the intersection between affective polarization and ideological extremity (those individuals who describe themselves as being on the far-left or far-right), evidence from 45 countries shows that the most ideologically extreme individuals are usually those who display the highest out-group affect, but that this relationship is stronger in older democracies, in countries with parliamentary systems, and where the party system of a country is less fragmented⁸⁰.

Beyond institutions, socio-economic factors and income inequality are often implicated as drivers of polarization^{47,73,81–86}. Evidence from 20 high-income democracies (Europe, USA, Canada, New Zealand, Australia) shows that affective polarization is worse when levels of income inequality and unemployment are high⁷³. But, again, countries with high income inequality show lower levels of ideological polarization amongst the mass public (e.g., the U.S.) than lower income inequality countries (e.g., Sweden^{81,86}). Other important institutional factors may alter how polarization manifests such as cross-national differences in political engagement⁷³ and the role of supra-national organisations (e.g., the European Union, or NATO⁸⁷). However, it is often hard to establish

causal evidence of a relationship. The difficulty of acquiring causal evidence for what is, and is not, polarizing is one of the major challenges facing researchers. And arguably, this issue is worsened by studying polarization in domain-silos, since such research cannot address the complex societal feedback loops involved in driving and sustaining the phenomenon.

Finally, it is often assumed that polarization has roots in ethnic or religious differences (Box 3). However, while these factors may correlate with observed differences in polarization, and may be used as the basis of polarizing strategies by political leaders, multiple comparative studies have shown that social cleavages do not, inherently, predict damaging forms of polarization^{60,88}. For further detail on the drivers of affective polarization and how it manifests across comparative contexts (including in the Global North, Latin America, and Asia) we point the reader to the recently published Handbook of Affective Polarization⁸⁹.

Polarization in the information age**The role of the news media**

Partisan news media are often assumed to be polarizing, but concrete evidence detailing the polarizing role of the news media is mixed and beset by technical issues and geographic limitations. Again, the available literature is heavily U.S. and Europe-centric⁹⁰ and the role of the news media in developing countries or in autocracies is rarely considered (a discussion of why this research is challenging is provided in Box 2).

Most studies on the polarizing role of the news media concern either the supply side of the news—how the news media ecosystem is aligned with specific political ideologies—or the reception side—namely what news individuals consume. Addressing the former, Western news media systems have been categorized according to “political parallelism”, referring to the one-to-one correspondence between news outlets and political parties^{91,92}. Using this framework, countries were classified as having low levels of political parallelism (e.g., the U.S., UK, Ireland), high levels but lowering over time (e.g., Austria, Belgium, Finland, Germany), or sustained high levels (e.g., Italy, France, Spain, Greece), revealing substantial differences in the politicization of the news media across developed countries, indicative of differences in ideological polarization.

With regard to the consumption side, the proliferation of the internet has enabled easy access to increasingly diverse online content, some of which is highly partisan or radical^{36,93,94}. Nevertheless, survey self-reports in 12 developed countries find little ideological bias in news consumption patterns²² with most individuals favouring centrist and moderate news sources, both on- and offline. This is despite fears that social media algorithms and “filter bubbles” will drive users to more radical content which in turn may drive increasing affective polarization (note that evidence for the existence of filter bubbles is very limited^{95–98}). The U.S. has the most partisan news consumption, and individual communities do have an ideologically skewed news media diet (more so in the U.S. than other countries^{28,99}), but absolute levels of partisan news consumption remain low. How to compare these results to countries where the politicization of the news media is more explicit is unclear, especially in authoritarian states where news outlets are forced to support the regime in power or risk political backlash¹⁰⁰.

But questions of news supply and demand do not answer whether the news media impact population-level polarization across the general population—do we expect changes in the news media to affect the polarization trends measured by traditional survey studies? Limited experimental evidence from the U.S. and Israel suggests that exposure to partisan news outlets and their messages reinforces prior political attitudes and affective polarization^{94,101–103}. However, given the challenges in measuring exposure to (partisan) media using self-reports and in ascertaining their effects outside experimental settings, this research may not portray the role of the news media in the real world accurately. In this context, a major challenge is acquiring high-quality causal evidence¹⁰⁴. This is because standard research designs are ill equipped to disentangle the effects of media exposure on the pre-existing beliefs and behaviours of experimental subjects, from the effect of experimental media treatments¹⁰⁵. Additionally, establishing a population-level baseline for the content individuals are exposed to in their day-to-day lives is far from trivial, and may be conflated by self-selection biases in how individuals report their media consumption habits¹⁰⁶.

Without high-quality causal evidence, disentangling whether partisan media is a reflection of underlying population-level polarization due to self-selection biases, or whether partisan media is the cause of polarization, remains difficult¹⁰⁵. Partly addressing this challenge, a 2020 study incentivised regular Fox News viewers in the U.S. to watch CNN for a month, finding that this switch had a moderating effect on their political attitudes¹⁰⁷, thereby implying that continued exposure to partisan media is polarizing. Yet, other work finds null effects from exposure to politically aligned news media: Online behavioral data shows that partisan news consumption has little impact on attitudes or affective polarization in the U.S.^{33,108,109}, but how these results translate to other world regions is unclear.

Finally, the global loss of trust in the news, and very low engagement with news media content in general, mean that some researchers (but certainly not all) are increasingly questioning the polarizing role of the news media^{110–113}. Self-surveys carried out across 46 countries (in Europe, the Americas, Asia, and some in Africa) show that interest in the news has fallen in the last decade¹⁹. No country has seen a growth in interest, and few countries have maintained stable levels—notably Finland, the Netherlands, and South Korea. Providing direct evidence of this, studies using web browsing data show that only 3.4% of all URLs visited by large samples in the U.S., Poland, and the Netherlands are news, and, even when individuals visit partisan domains, more than half of the content they consume is apolitical, such as sports or food recipes¹⁰⁸; see also ref. 114.

On social media—which has become a leading gateway for accessing news content¹⁹—news makes up only a small part of a user’s feed (less than 3% for U.S. Facebook users¹¹⁵). This lack of interest in the news and political content may partly explain diverging trends in mass and elite polarization across countries. However, a competing

explanation suggests that the lack of engagement with highly politicized hot-button issues is not the key for understanding increasing affective polarization. Rather, a broader set of beliefs and preferences have become captured in a network of ideologically aligned belief packages (what DellaPosta refers to as an “oil spill”⁶⁷), capturing issues which are seemingly apolitical (e.g., liberals’ preference for Lattes¹¹⁶), such that over time fewer issue-preferences remain which cut across partisan lines. Empirical research suggests that digital media play a key role in facilitating this process, whereby “identities, beliefs, and cultural preferences become drawn into an all-encompassing societal division”⁶⁸, but evidence from outside the U.S. is limited.

Rather than reducing exposure to political content, some argue that increasing public exposure to high-quality news may have a depolarizing effect. This would pull more moderate citizens into the democratic process, increase public resilience to misinformation^{117–119}, and decrease their susceptibility to populist or hyper-partisan rhetoric which is believed to play a key role in driving affective polarization³⁶. However, the widespread decline in trust in the news (16 out of 19 surveyed countries¹⁹) casts a shadow on this approach (see also¹¹² for a discussion on the distinction between a lack of trust, and mistrust, of the media). Strong public service broadcasters may play a key role in maintaining trust, but in many markets politicians and other public figures are actively undermining trust through their continued criticism of the news media, e.g., in Greece, Peru, the U.S., and the UK¹⁹, or have deliberately defunded public service broadcasting.

The role of the internet and social media

The internet and social media are playing an increasingly dominant role in how citizens (and politicians) produce and consume content, but this fundamental change to our lives is too often ignored in traditional polarization research. This should not be understood to imply that the internet and social media are ultimately responsible for recent trends in polarization—claims of social media’s negative impacts are often anecdotal, or based on correlational evidence¹²⁰, and there are substantial differences across countries in the extent to which social media is seen as democratically harmful (Americans see social media as far more damaging than other surveyed populations¹²¹)—but it is critical to understand the role that this new information ecosystem may play in driving polarization.

With a focus on user-user interactions, one perspective argues that social media drives polarization (both ideological and affective) by enabling politically engaged users to primarily interact with other ideologically aligned individuals. This phenomenon is sometimes referred to as interactional polarization^{9,122}, and is observed across many platforms (albeit to varying extents^{123–125}) and countries (e.g., in the U.S.^{126,127}, Brazil¹²⁸, Egypt¹²⁹, Israel¹²², Venezuela¹³⁰, Pakistan¹³¹, several Western countries⁹, and Japan¹³²). These patterns are considered to be indicative of an epistemic bubble, or echo chamber^{123,133,134}, driving greater ideological isolation, limiting out-group exposure, and thereby generating greater out-group affect¹³⁵.

However, another perspective argues that these homophilous interaction patterns do not constitute ideological isolation. First, social media interactions in most countries are not primarily with political actors or content^{9,127,136,137}, and the interaction patterns of politically engaged users differ to those who are less politically engaged⁹. Second, when individuals interact with news content rather than individuals, their news diet is often politically moderate^{22,99,109}. Third, while certain interaction types (like reposts) give the appearance of isolated communities, other interaction types (mentions or follows) reveal that out-group interaction is not so rare^{9,134,138,139}, albeit these interactions are often toxic attacks (a feature common across many countries^{9,140}). These differences across interaction types also translate to differences across user groups, with more active users less likely to view social media content as toxic and unconstructive, when compared to passive users (“lurkers”)¹⁴¹. Fourth, empirical evidence suggests that, contrary

to the echo chamber hypothesis, ideologically isolated interactions on social media may increase the plurality of cross-cutting beliefs across user groups, whereas non-local interactions across partisan lines are precisely those which drive the alignment of ideology and identity⁶⁸. Finally, even if online interactions are ideologically homogeneous, individuals may in fact be more isolated in their offline environments^{82,142,143}, a factor likely to be highly country- and context-dependent. For these reasons, some scholars have suggested that research should look beyond the mere act of interaction and instead focus on interaction quality as well as on promoting prosocial behavior^{144–147}.

Despite these ongoing debates, we must not overlook the fact that social media has become a cornerstone of contemporary political campaigns, and many are convinced that effective social media campaigns have the power to swing political elections^{148,149}, with the 2016 election of Donald Trump to the U.S. presidency being one key example. Evidence of substantial user heterogeneity in the engagement with, and sharing of, political content on social media¹⁵⁰ suggests that there are likely substantial differences in the polarizing effect of social media across user groups, possibly related to factors such as online literacy (which we note is hard to measure accurately¹⁵¹). However, given that elections are often won on fine margins, the persuasive effects of social media on small subpopulations may still have substantial effects across broader society.

These biases in exposure also apply to political elites. When strategists for the UK Labour Party were asked in 2022 how to improve the party's electability, they cited MPs' use of Twitter as one of the biggest impediments to winning: "For members of parliament, Twitter has become a constant focus group on what (allegedly) matters"¹⁵². This is crucial—we must study the polarizing role of social media not because it is necessarily representative of the wider population, but precisely because its unrepresentativeness may skew perceptions of public and elite opinion.

A core question in the study of digital media and polarization is whether biased information exposure leads groups to adopt more extreme views, driving changes in the level of polarization measured in political values surveys. The largest Facebook field experiment to date—a collaboration with Meta which was limited to U.S. researchers³⁵—finds that reducing users' exposure to content from like-minded sources does not measurably affect affective polarization or ideological extremity²⁷. However, similar research (with smaller study sizes) suggests that such results are likely to be country-dependent: A U.S. study found that deactivating Facebook reduced both ideological and affective polarization¹⁵³, while a replication in Bosnia-Herzegovina found an increase in polarization¹⁵⁴, with evidence to suggest that this unexpected increase may be related to individuals' lack of offline exposure to other ethnicities (although this result did not replicate in another ethnically polarized setting, Cyprus¹⁵⁵).

A key area of differentiation across countries is the different social platforms which dominate political messaging. In the last decade, Twitter, now X, and Facebook have been the dominant platforms for political communication in high-income democracies. But in countries such as Brazil, the Philippines, and India, WhatsApp is a substantially more important channel than X or Facebook for the dissemination of political content from political parties to voters¹⁵⁶. In Eastern European states, Telegram plays a similar role in facilitating political communication. These comparative differences are important because platforms based around private peer-to-peer, or group, messaging pose a major barrier to data acquisition for academic research.

The risks of AI and bots exacerbating polarization

There are broad concerns that the polarizing effect of social media may be exacerbated by recommendation algorithms, AI generated content, and bots^{36,157–159}, especially given the opaqueness of how these

systems operate, and the scales at which they can affect the information that individuals are exposed to.

Since the quantity of online content vastly exceeds what any individual can consume, social media companies must—by necessity—curate the content which their users are exposed to. For example, on Youtube recommended content makes up over 70% of the videos watched¹⁶⁰. However, with AI-based recommendation systems optimizing for user engagement¹⁶¹, there are fears of a self-reinforcing cycle in which consumers are exposed to increasingly homogeneous content which—if the content is political in nature—may drive polarization (although this is not inevitable; AI could also be used to counter polarization, see below).

Unfortunately, the commercial interests of tech companies mean that independently auditing this polarizing role is difficult. From the few studies available, U.S. data on Instagram and Facebook suggest that various changes to their recommendation algorithms do not impact ideological or affective polarization. These changes include removing re-shared content³³, removing all content rankings³⁴, and down-ranking content from ideologically aligned sources²⁷. In contrast, recent work, again U.S. focused, shows that changes to algorithmic ranking on X can both raise or lower affective polarization depending on how content expressing partisan animosity, or anti-democratic attitudes, was ranked²⁹. On Youtube, data suggests that recommendations are largely reflective of individual preferences, and that the algorithm only has a mild role in driving users towards more moderate political content³⁷.

Similar fears are raised concerning how AI generated content (often shared by bots) alters social media. Many studies have shown that bots amplify political content^{157,159,162,163}, and a growing body of evidence demonstrates how large language models (LLMs) are often ideologically biased¹⁶⁴, and can be tuned to promote specific political positions (e.g., xAI's Grok, which was intended to counteract the "woke AI" bias of other models¹⁶⁵). These observations have driven concerns that swarms of AI bots may, for example, produce a fabricated grassroots consensus, a fragmented shared reality, suppress or mobilise voters at the micro level, and erode institutional trust¹⁶⁶.

A global survey covering the elections of 2024 suggests that initial fears about the negative democratic impacts of generative AI were overblown¹⁶⁷, but there is still cause for concern. One major study shows that LLMs are highly effective when deployed for political persuasion¹⁶⁸, especially if they have been fine-tuned for persuasiveness. However, this capacity for persuasion can also be leveraged for good. Multiple studies demonstrate how LLMs can be used for democratic deliberation, to correct misconceptions, and to durably reduce conspiracy beliefs^{146,169–171}. Crucially, whether AI will drive polarization, or contribute to mitigating its damaging consequences, is fundamentally a question of how AI is used and the different cultural contexts in which it is used, not an intrinsic property of the technology itself.

Understanding the polarizing role of AI in non-experimental settings requires the ability to identify AI and bot generated content using appropriate detection tools. But the efficacy of these tools is context, and therefore country, dependent. For example, recent research shows that Google's toxicity detection tool "Perspective" is far more likely to label content in German as toxic, than content in other languages¹⁷². These challenges may be exacerbated as generative AI further improves and it becomes harder to distinguish between machine- and user-generated content¹⁷³. Improved detection tools will be required to distinguish between the true population-level impact of AI from a mirage of false polarization which reflects the actions of bots rather than human actors¹⁶⁶. Such tools are especially necessary for low-resource languages. However, researchers should be conscious that the development of such tools can be a double-edged sword if used for malicious purposes (e.g., in authoritarian contexts; see Box 2).

For a comprehensive discussion on the possible polarizing role of AI, especially if exploited by bad actors, see ref. 174.

Limiting the internet's polarizing role

If we are to prevent digital media from driving polarization, we must consider how social media and AI companies moderate their platforms, and the country-specific regulatory environments in which they operate^{175,176}. On social media, state regulators can issue takedown requests to remove content they deem to be problematic at a national level (see for example, <https://help.x.com/en/rules-and-policies/post-withheld-by-country>). However, in most cases regulators cannot remove this content without using the—typically U.S.-based—social media companies as an intermediary. Furthermore, there have been numerous cases where takedown requests have been rejected, e.g., in India¹⁷⁷. This state of affairs puts an outsized emphasis on U.S. regulators, U.S. political interests, and a U.S. philosophy on platform regulation.

This outsized emphasis on U.S. interests has directly impacted the quality of polarization research. Notably, Meta's U.S. election studies—the gold standard for acquiring causal evidence into the online drivers of polarization^{27,33,34}—are part of a recent academic-industry collaboration. But such collaborations are currently dependent on the goodwill of, and the parameters set by, social media companies. These companies have incentives to prioritize U.S. research, which they pursue by granting researchers some “independence by permission”³⁵. Consequently, the Meta studies only provide evidence for the U.S., and only involve U.S.-based researchers. Other scientists may use more creative methods to causally estimate the role of social media in driving polarization^{37,178}, but without the ability to alter a social platform's algorithms, and with no passive consumption data, the questions which such studies can answer are limited.

These systemic biases are a barrier to diversifying polarization research leading some to call for social media research to be democratized^{30,31}. The European Union's new Digital Services Act (DSA) has the potential to transform access to online platform data, enabling a new phase of in-depth research, but only if the data access provided is sufficient for reliable and consistent measurements of online social phenomena (social media studies are especially prone to sampling bias¹³⁷). Early evidence suggests poor compliance with the DSA by tech companies³², and for researchers outside Europe and the U.S. barriers to data access will remain. Similarly, the EU's AI Act aims to establish a common regulatory and legal framework for the use of AI, and in doing so aims to limit AI's potentially harmful impacts, but whether this framework will be effective is unclear (in July 2025, Meta stated that it will not sign the EU's AI code of practice).

When considering digital media regulation at a global scale, it is important to emphasise that regulators (and governments) may make decisions in the interest of preventing online harms, but they could equally decide to manipulate online discourses in order to support the political regime in power^{179–181}. Either way, these choices are fundamentally political and rest on value judgments which differ between countries. Given the outsized role of U.S. platforms and regulators, there has been a growing desire for digital sovereignty¹⁸², especially with regards to the development of sovereign AI¹⁸³ and the regulation of online harms (e.g., Australia's social media ban for minors). Academics should be conscious of how diverging regulatory environments may bias future polarization research, and should consider how these environments may impact evolving polarization trends.

Challenges moving forward: countering polarization globally

Preventing the potential damaging societal impacts of political polarization requires effective and long-lasting interventions. However, the current dominance of U.S. research has put an outsized emphasis on developing interventions for affective polarization between

Democrats and Republicans in the U.S.²⁶, with few interventions having been tested elsewhere. These include measures promoting intergroup contact¹⁸⁴, emphasizing a shared identity between partisans¹⁸⁵, meaningful, non-political conversations across party lines^{186,187}, and inoculating against polarizing content^{188,189}. However, existing approaches often suffer from not being both durable (lasting for extended periods) and scalable (easily and effectively distributed)²⁵, and in some cases interventions counter-intuitively increase polarization, or result in other negative side-effects (see Box 1 in ref. 25; also^{154,190,191}). Relying on such narrowly designed measures risks introducing inappropriate or even counterproductive solutions in diverse societal and cultural contexts.

Fortunately, scholars are increasingly aware of the need to study interventions globally, with several projects already underway (for example globalsocialmediastudy.com/ and templetonworldcharity.org/LLPW). However, the impact of these studies risks remaining limited without systemic and cultural changes in how polarization research is produced and communicated.

Researchers working on understudied regions often face career disadvantages due to the perceived lack of visibility, relevance, and therefore impact, of their work^{192,193}. Academic institutions, publishers, and policymakers must actively promote and reward work that brings global balance to polarization research. This is not a call for less research on the U.S., which will likely remain prolific even if this research has been heavily impacted by recent funding cuts in the U.S., but rather to increase recognition and support for research that expands our understanding to include diverse global perspectives.

A first step to improving the current state of affairs, given the constraints on academic resources, is to prioritize comparative research that identifies both shared and unique aspects of polarization across different countries. As part of this, it is crucial that country-specific research avoids over-generalizing their results. For example, journals could mandate that studies focusing solely on the U.S. explicitly state this in their title or abstract (many currently don't), as is commonplace for research focusing on almost any other country. Additionally, particular efforts are needed to address the conceptual challenges of studying polarization in authoritarian states, and in low-income democracies where key political concepts, used widely in the Global North, may not apply (e.g., the framing of ideological polarization along a left-right axis^{15,51}).

To prevent further global disparities in polarization research, we argue that researchers must push back against growing restrictions on data access now. Since 2023, several social media companies (including Meta, X, Reddit) have shuttered academic access to their platforms' data, and new opportunities for data access such as the European Union's Digital Services Act remain largely unproven^{30–32}. If effective, these data provisions may enable continued observational research, but this will not enable the high-quality experimental research which is necessary to establish clear causal relationships between various aspects of polarization and digital media use^{37,178}. Such research will likely only be possible across countries, and at scale, with the buy-in of social media companies, or with substantially greater cooperation across the research community. Interdisciplinary, collaborative research is necessary to disentangle the complex web of factors driving different forms of polarization, and greater effort is required to establish the aspects of polarization on which there is academic consensus. One option here would be for interdisciplinary researchers to collaborate on the production of polarization “consensus statements”, as is common in medical research.

Building a globally equitable picture of political polarization, and its complex manifestations and drivers across different publics and media modalities, requires integrating research methods from the social sciences, with the advanced data science approaches which are necessary for understanding the structure of, and spread of information through, our digital world, and how this content impacts the

political beliefs of individuals. Examples of promising methods which bridge disciplinary divides include study designs leveraging matched survey and digital trace data⁵⁸, including data acquired through data donation frameworks⁵⁹, and novel computational methods for understanding the interdependent structure of political attitudes using networks^{194,195}. However, for researchers newly entering the study of political polarization, there is a burden of responsibility to ensure that proposed technical solutions to key research challenges remain conscious of polarization's conceptual nuances. Complex feedback loops between different aspects of polarization cannot be understood if this nuance is lost.

Finally, as our understanding of polarization evolves and new interventions are developed, both academics and policymakers must account for the unique regulatory and practical challenges of each country^{176,196}. Interventions must be both ethically and practically tailored to the local context. We cannot assume that societies outside the U.S.—where empirical evidence is comparably limited—face the same problems which can be solved with the same solutions. Academics play a critical role in gathering evidence to better understand how polarization manifests across countries, but the implementation of interventions is necessarily political, and often controversial. These decisions must rest with local stakeholders and decision makers. However, without high-quality evidence which is contextually appropriate, even the best-intentioned decision makers may draw inappropriate conclusions from flawed evidence, risking the further erosion of democratic institutions.

References

- McCoy, J., Rahman, T. & Somer, M. "Polarization and the global crisis of democracy: common patterns, dynamics, and pernicious consequences for democratic polities". *Am. Behav. Sci.* **62**, 16–42 (2018).
- World Economic Forum, "The Global Risks Report 2025." <https://www.weforum.org/publications/global-risks-report-2025/in-full/> (2025).
- IPSOS, "Global Trends about Polarization: A Divided World," Global Trends about Polarization: A Divided World. <https://www.ipsos.com/sites/default/files/2023-Ipsos-Global-Trends-Report.pdf> (2023).
- Norris, P. & Inglehart, R. *Cultural backlash: Trump, Brexit, and authoritarian populism*. Cambridge University Press, (2019).
- Orhan, Y. E. The relationship between affective polarization and democratic backsliding: comparative evidence. *Democratization* **29**, 714–735 (2022).
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S. J. The origins and consequences of affective polarization in the United States. *Annu. Rev. Polit. Sci.* **22**, 129–146 (2019).
- Wagner, M. Affective polarization in multiparty systems. *Elect. Stud.* **69**, 102199 (2021).
- Gidron, N., Adams, J. & Horne, W. *American Affective Polarization In Comparative Perspective*. Cambridge University Press (2020).
- Falkenberg, M., Zollo, F., Quattrociocchi, W., Pfeffer, J. & Baronchelli, A. Patterns of partisan toxicity and engagement reveal the common structure of online political communication across countries. *Nat. Commun.* **15**, 9560 (2024).
- Citrin, J. & Stoker, L. Political trust in a cynical age. *Annu. Rev. Polit. Sci.* **21**, 49–70 (2018).
- Hetherington, M. J. & Rudolph, T. J. *Why Washington Won't Work: Polarization, Political Trust, and the Governing Crisis*, Illustrated edition. University of Chicago Press, (2015).
- Pew Research Center. As Partisan Hostility Grows, Signs of Frustration With the Two-Party System. https://www.pewresearch.org/wp-content/uploads/sites/20/2022/08/PP_2022.09.08_partisan-hostility_REPORT.pdf (2022).
- Pew Research Center, "The polarization in today's Congress has roots that go back decades,". <https://www.pewresearch.org/short-reads/2022/03/10/the-polarization-in-todays-congress-has-roots-that-go-back-decades/> (2024).
- Kalmoe, N. P. & Mason, L. *Radical American partisanship: mapping violent hostility, its causes, and the consequences for democracy*. in *Chicago studies in American politics*. The University of Chicago Press, (2022).
- Nielsen, L. H. & Petersen, M. B. Beyond left and right: the role of system trust in COVID-19 attitudes and behaviors across eight western countries. *Eur. J. Polit. Res.* **64**, 1233–1256 (2025).
- Falkenberg, M. et al. Growing polarization around climate change on social media. *Nat. Clim. Change* **12**, 1114–1121 (2022).
- Caldwell, D., Cohen, G. & Vivyan, N. Long-run trends in partisan polarization of climate policy-relevant attitudes across countries. *Environ. Polit.* <https://doi.org/10.1080/09644016.2024.2403957> (2024).
- McCoy, J., Press, B., Somer, M., & Tuncel, O. "Reducing Pernicious Polarization: A Comparative Historical Analysis of Depolarization," Carnegie Endowment for International Peace, https://carnegie-production-assets.s3.amazonaws.com/static/files/McCoy_et_al_-_Polarization_final_3.pdf (2022).
- Newman, N., Fletcher, R., Eddy, K., Robertson, C. & Nielsen, R. "Digital news report 2023," Reuters Institute for the Study of Journalism, (2023).
- Röllicke, L. Polarisation, identity and affect - conceptualising affective polarisation in multi-party systems. *Elect. Stud.* **85**, 102655 (2023).
- Carothers, T. & O'Donoghue, A. "Introduction," in *Democracies Divided*, Carothers, T. & O'Donoghue, A., Eds., in *The Global Challenge of Political Polarization*, Brookings Institution Press, (2019).
- Fletcher, R., Cornia, A. & Nielsen, R. K. How polarized are online and offline news audiences? a comparative analysis of twelve countries. *Int. J. Press.* **25**, 169–195 (2020).
- Finkel, E. J. et al. Political sectarianism in America. *Science* **370**, 533–536 (2020).
- Jost, J. T., Baldassarri, D. S. & Druckman, J. N. Cognitive-motivational mechanisms of political polarization in social-communicative contexts. *Nat. Rev. Psychol.* **1**, 560–576 (2022).
- Hartman, R. et al. Interventions to reduce partisan animosity. *Nat. Hum. Behav.* **6**, 1194–1205 (2022).
- Voelkel, J. G. et al. Megastudy testing 25 treatments to reduce antidemocratic attitudes and partisan animosity. *Science* **386**, eadh4764 (2024).
- Nyhan, B. et al. Like-minded sources on Facebook are prevalent but not polarizing. *Nature* **620**, 137–144 (2023).
- González-Bailón, S. et al. Asymmetric ideological segregation in exposure to political news on Facebook. *Science* **381**, 392–398 (2023).
- Piccardi, T. et al. Reranking partisan animosity in algorithmic social media feeds alters affective polarization. *Science* **390**, eadu5584 (2025).
- Roozenbeek, J. & Zollo, F. Democratize social-media research—with access and funding. *Nature* **612**, 404–404 (2022).
- Rathje, S. To tackle social-media harms, mandate data access for researchers. *Nature* **633**, 36–36 (2024).
- Jaursch, J., Ohme, J. & Klinger, U. "Enabling research with publicly accessible platform data: early DSA compliance issues and suggestions for improvement," <https://doi.org/10.34669/WI.WPP/9> (2024).
- Guess, A. M. et al. Reshares on social media amplify political news but do not detectably affect beliefs or opinions. *Science* **381**, 404–408 (2023).

34. Guess, A. M. et al. How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science* **381**, 398–404 (2023).
35. Wagner, M. W. Independence by permission. *Science* **381**, 388–391 (2023).
36. Haroon, M. et al. Auditing YouTube’s recommendation system for ideologically congenial, extreme, and problematic recommendations. *Proc. Natl. Acad. Sci.* **120**, e2213020120 (2023).
37. Hosseinmardi, H. et al. Causally estimating the effect of YouTube’s recommender system using counterfactual bots. *Proc. Natl. Acad. Sci.* **121**, e2313377121 (2024).
38. Santos, F. P., Lelkes, Y. & Levin, S. A. Link recommendation algorithms and dynamics of polarization in online social networks. *Proc. Natl. Acad. Sci.* **118**, e2102141118 (2021).
39. Abramowitz, A. I. & Saunders, K. L. Is polarization a myth? *J. Polit.* **70**, 542–555 (2008).
40. Fiorina, M. P. & Abrams, S. J. Political polarization in the American public. *Annu. Rev. Polit. Sci.* **11**, 563–588 (2008).
41. DiMaggio, P., Evans, J. & Bryson, B. Have American’s social attitudes become more polarized? *Am. J. Sociol.* **102**, 690–755 (1996).
42. Sanders, D. et al. *Citizens and the European Polity: Mass Attitudes Towards The European And National Politics, First Edition*. In *Intune Series*. Oxford: Oxford University Press, (2012).
43. Ford, R. & Jennings, W. The changing cleavage politics of Western Europe. *Annu. Rev. Polit. Sci.* **23**, 295–314 (2020).
44. Bramson, A. et al. Understanding polarization: meanings, measures, and model evaluation. *Philos. Sci.* **84**, 115–159 (2017).
45. Iyengar, S., Sood, G. & Lelkes, Y. Affect, not ideology. *Public Opin. Q.* **76**, 405–431 (2012).
46. Mason, L. et al. I disrespectfully agree’: the differential effects of partisan sorting on social and issue polarization: partisan sorting and polarization. *Am. J. Polit. Sci.* **59**, 128–145 (2015).
47. Dalton, R. J. Modeling ideological polarization in democratic party systems. *Elect. Stud.* **72**, 102346 (2021).
48. Emanuele, V. & Marino, B. Party system ideological polarization in Western Europe: data, trends, drivers, and links with other key party system properties (1945–2021). *Polit. Res. Exch.* **6**, 2399095 (2024).
49. Young, D. J. et al. A new measure of issue polarization using *k*-means clustering: US trends 1988–2024 and predictors of polarization across the world. *R. Soc. Open Sci.* **13**, 251428 (2026).
50. Pew Research Center, “Political Polarization in the American Public,” <https://www.pewresearch.org/wp-content/uploads/sites/4/2014/06/6-12-2014-Political-Polarization-Release.pdf> (2014).
51. Geser, H. The limits of ideological globalization. Current patterns of ‘left and right’ in different geographical regions. <https://doi.org/10.5167/UZH-48431> (2008).
52. Beattie, P., Chen, R. & Bettache, K. When left is right and right is left: the psychological correlates of political ideology in China. *Polit. Psychol.* **43**, 457–488 (2022).
53. Pan, J. & Xu, Y. China’s Ideological Spectrum. *J. Polit.* **80**, 254–273 (2018).
54. Norris, P. et al. Measuring populism worldwide. *Party Polit* **26**, 697–717 (2020).
55. Boxell, L., Gentzkow, M. & Shapiro, J. M. “Cross-country trends in affective polarization.” *Rev. Econ. Stat.* https://doi.org/10.1162/rest_a_01160 (2022).
56. Reiljan, A., Garzia, D., Ferreira Da Silva, F. & Trechsel, A. H. Patterns of affective polarization toward parties and leaders across the democratic world. *Am. Polit. Sci. Rev.* **118**, 654–670 (2024).
57. Coppedge, M. et al. *Varieties of Democracy: Measuring Two Centuries of Political Change*. (Cambridge University Press, 2020).
58. Stier, S., Breuer, J., Siegers, P. & Thorson, K. Integrating survey data and digital trace data: key issues in developing an emerging field. *Soc. Sci. Comput. Rev.* **38**, 503–516 (2020).
59. Kmetty, Z. et al. Determinants of willingness to donate data from social media platforms. *Inf. Commun. Soc.* **28**, 1324–1349 (2025).
60. McCoy, J. & Somer, M. Toward a theory of pernicious polarization and how it harms democracies: comparative evidence and possible remedies. *Ann. Am. Acad. Pol. Soc. Sci.* **681**, 234–271 (2019).
61. Kreiss, D. & McGregor, S. C. A review and provocation: on polarization and platforms. *New Media Soc* **26**, 556–579 (2024).
62. Rogowski, J. C. & Sutherland, J. L. How Ideology Fuels Affective Polarization. *Polit. Behav.* **38**, 485–508 (2016).
63. Webster, S. W. & Abramowitz, A. I. The ideological foundations of affective polarization in the U.S. electorate. *Am. Polit. Res.* **45**, 621–647 (2017).
64. Mason, L. et al. *Uncivil agreement: How Politics Became Our Identity*. (University of Chicago Press, 2018).
65. Brown, J. R. & Enos, R. D. The measurement of partisan sorting for 180 million voters. *Nat. Hum. Behav.* **5**, 998–1008 (2021).
66. Lelkes, Y. et al. Affective polarization and ideological sorting: a reciprocal, albeit weak, relationship. *Forum (Genova)* **16**, 67–79 (2018).
67. DellaPosta, D. et al. Pluralistic collapse: the ‘oil spill’ model of mass opinion polarization. *Am. Sociol. Rev.* **85**, 507–536 (2020).
68. Törnberg, P. et al. How digital media drive affective polarization through partisan sorting. *Proc. Natl. Acad. Sci.* **119**, e2207159119 (2022).
69. Schul, Y. & Ganzach, Y. Divergent views of party positions: How ideology and own issue position shape party perception through convergence and divergence processes. *Group Process. Inter-group Relat.* **28**, 256–276 (2025).
70. Bankert, A., Huddy, L. & Rosema, M. Measuring partisanship as a social identity in multi-party systems. *Polit. Behav.* **39**, 103–132 (2017).
71. Brewer, M. B. & Pierce, K. P. Social identity complexity and out-group tolerance. *Pers. Soc. Psychol. Bull.* **31**, 428–437 (2005).
72. Brewer, M. B. Multiple identities and identity transition: implications for Hong Kong. *Int. J. Intercult. Relat.* **23**, 187–197 (1999).
73. Gidron, N., Adams, J. & Horne, W. “How ideology, economics and institutions shape affective polarization in democratic polities,” in *Annual conference of the American Political Science Association*, (2018).
74. Reiljan, A. et al. Fear and loathing across party lines’ (also) in Europe: affective polarisation in European party systems. *Eur. J. Polit. Res.* **59**, 376–396 (2020).
75. The American Political Science Review, “Part III. The Prospect for Action,” *Am. Polit. Sci. Rev.* **44**, 85, <https://doi.org/10.2307/1951001> (1950).
76. Hobolt S. B. & Tilley, J. *Affective Polarization Around Issues*. in *Handbook of Affective Polarization*, (eds. Torcal, M. & Harteveld, E.) <https://doi.org/10.4337/9781035310609.00025> (Edward Elgar Publishing, 2025).
77. Young, D. J. & de-Wit, L. H. Affective polarization within parties. *Polit. Psychol.* <https://doi.org/10.1111/pops.12973>. (2024).
78. Areal, J. Them’ without ‘us’: negative identities and affective polarization in Brazil. *Polit. Res. Exch.* **4**, 2117635 (2022).
79. Vowles, J., Banducci, S. A. & Karp, J. A. Forecasting and evaluating the consequences of electoral change in New Zealand. *Acta Polit* **41**, 267–284 (2006).
80. Riera, P. & Madariaga, A. G. Overlapping polarization: on the contextual determinants of the interplay between ideological and affective polarization. *Elect. Stud.* **84**, 102628 (2023).
81. Iversen, T. & Soskice, D. Information, inequality, and mass polarization: ideology in advanced democracies. *Comp. Polit. Stud.* **48**, 1781–1813 (2015).

82. Monti, C., D'Ignazi, J., Starnini, M. & De Francisci Morales, G. "Evidence of Demographic rather than Ideological Segregation in News Discussion on Reddit," in *Proceedings of the ACM Web Conference 2023*, <https://doi.org/10.1145/3543507.3583468> (2023).
83. Stewart, A. J., McCarty, N. & Bryson, J. J. Polarization under rising inequality and economic decline. *Sci. Adv.* **6**, eabd4201 (2020).
84. McCarty, N. M. et al. *Polarisation: What Everyone Needs To Know*. (University Press, 2019).
85. N. M. McCarty, K. T. Poole, and H. Rosenthal, *Polarized America: the dance of ideology and unequal riches*, Second edition. in *The Walras-Pareto Lectures*. The MIT Press (2016).
86. Hübscher, E., Sattler, T. & Wagner, M. Does austerity cause polarization? *Br. J. Polit. Sci.* **53**, 1170–1188 (2023).
87. Xia, Y. et al. "The Russian invasion of Ukraine selectively depolarized the Finnish NATO discussion," Preprint arXiv: <http://arxiv.org/abs/2212.07861> (2023).
88. Somer, M. & McCoy, J. Déjà vu? polarization and endangered democracies in the 21st century. *Am. Behav. Sci.* **62**, 3–15 (2018).
89. M. Torcal and E. Hartevelde, Eds., *Handbook of affective polarization*. in *Elgar handbooks in political science*. Cheltenham, UK; Northampton, MA, USA: Edward Elgar Publishing, (2025).
90. Kubin, E. & Von Sikorski, C. The role of (social) media in political polarization: a systematic review. *Ann. Int. Commun. Assoc.* **45**, 188–206 (2021).
91. D. C. Hallin and P. Mancini, *Comparing Media Systems: Three Models Of Media And Politics*. Cambridge University Press, (2004).
92. Brüggemann, M., Engesser, S., Büchel, F., Humprecht, E. & Castro, L. Hallin and Mancini revisited: four empirical types of western media systems: hallin and mancini revisited. *J. Commun.* **64**, 1037–1065 (2014).
93. Van Aelst, P. et al. Political communication in a high-choice media environment: a challenge for democracy? *Ann. Int. Commun. Assoc.* **41**, 3–27 (2017).
94. Stroud, N. J. et al. Polarization and partisan selective exposure. *J. Commun.* **60**, 556–576 (2010).
95. Bruns, A. et al. *Are filter bubbles real? in Digital futures*. (Cambridge, UK; Medford, MA: Polity Press, 2019).
96. Ross Arguedas, A., Robertson, C. T., Fletcher, R. & Nielsen, R. K. "Echo chambers, filter bubbles, and polarisation: a literature review," *Reuters Institute for the Study of Journalism*. <https://doi.org/10.60625/RISJ-ETXJ-7K60> (2022).
97. Cardenal, A. S., Aguilar-Paredes, C., Galais, C. & Pérez-Montoro, M. Digital technologies and selective exposure: how choice and filter bubbles shape news media exposure. *Int. J. Press.* **24**, 465–486 (2019).
98. Robertson, R. E. et al. Users choose to engage with more partisan news than they are exposed to on Google search. *Nature* **618**, 342–348 (2023).
99. Mangold, F., Schoch, D. & Stier, S. Ideological self-selection in online news exposure: evidence from Europe and the US. *Sci. Adv.* **10**, eadg9287 (2024).
100. Enikolopov R. & Petrova, M. *Media Capture: Empirical Evidence*. in *Handbook of Media Economics*, vol. 1, 687–700 (Elsevier, 2015).
101. Garrett, R. K. et al. Implications of pro- and counterattitudinal information exposure for affective polarization: partisan media exposure and affective polarization. *Hum. Commun. Res.* **40**, 309–332 (2014).
102. Garrett, R. K., Long, J. A. & Jeong, M. S. From partisan media to misperception: affective polarization as mediator. *J. Commun.* **69**, 490–512 (2019).
103. Levendusky, M. S. Why do partisan media polarize viewers? *Am. J. Polit. Sci.* **57**, 611–623 (2013).
104. Prior, M. et al. Media and political polarization. *Annu. Rev. Polit. Sci.* **16**, 101–127 (2013).
105. J. De Benedictis-Kessner, M. A. Baum, and A. J. Berinsky, "Polarization and Media Usage: Disentangling Causality," in *The Oxford Handbook of Electoral Persuasion*, E. Suhay, B. Grofman, and A. H. Trechsel, Eds., Oxford University Press, 553–571 <https://doi.org/10.1093/oxfordhb/9780190860806.013.22> (2020).
106. Arendt, F. Media stereotypes, prejudice, and preference-based reinforcement: toward the dynamic of self-reinforcing effects by integrating audience selectivity. *J. Commun.* **73**, 463–475 (2023).
107. Broockman, D. E. & Kalla, J. L. Consuming cross-cutting media causes learning and moderates attitudes: A field experiment with Fox News viewers. *J. Polit.* **87**, 246–261 (2025).
108. Wojcieszak, M. et al. "No polarization from partisan news: over-time evidence from trace data,". *Int. J. Press.* **28**, 601–626 (2023).
109. Guess, A. M. et al. (Almost) everything in moderation: new evidence on Americans' online media diets. *Am. J. Polit. Sci.* **65**, 1007–1022 (2021).
110. Strömbäck, J. et al. News media trust and its impact on media use: toward a framework for future research. *Ann. Int. Commun. Assoc.* **44**, 139–156 (2020).
111. Fletcher, R. et al. The link between changing news use and trust: longitudinal analysis of 46 countries. *J. Commun.* **75**, 1–15 (2025).
112. Park, S. et al. The relationship between news trust, mistrust and audience disengagement. *Journalism* **26**, 2285–2304 (2025).
113. Guess, A. M., Barberá, P., Munzert, S. & Yang, J. The consequences of online partisan media. *Proc. Natl. Acad. Sci.* **118**, e2013464118 (2021).
114. Allen, J., Howland, B., Mobius, M., Rothschild, D. & Watts, D. J. Evaluating the fake news problem at the scale of the information ecosystem. *Sci. Adv.* **6**, eaay3539 (2020).
115. Eisenach, J. A. et al. Meta and the News: Assessing the Value of the Bargain. NERA Economic Consulting, <https://about.fb.com/wp-content/uploads/2023/03/NERA-Global-White-Paper-FINAL-March-2023.pdf> (2023).
116. DellaPosta, D., Shi, Y. & Macy, M. Why Do Liberals Drink Lattes? *Am. J. Sociol.* **120**, 1473–1511 (2015).
117. Arechar, A. A. et al. Understanding and combatting misinformation across 16 countries on six continents. *Nat. Hum. Behav.* **7**, 1502–1513 (2023).
118. Del Vicario, M. et al. The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **113**, 554–559 (2016).
119. Van Der Linden, S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nat. Med.* **28**, 460–467 (2022).
120. Lorenz-Spreen, P., Oswald, L., Lewandowsky, S. & Hertwig, R. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nat. Hum. Behav.* **7**, 74–101 (2022).
121. Pew Research Center, "Social media seen as mostly good for democracy across many nations, but U.S. is a major outlier," (2022).
122. Yarchi, M., Baden, C. & Kligler-Vilenchik, N. Political polarization on the digital sphere: a cross-platform, over-time analysis of inter-actional, positional, and affective polarization on social media. *Polit. Commun.* **38**, 98–139 (2021).
123. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. & Starnini, M. The echo chamber effect on social media. *Proc. Natl. Acad. Sci.* **118**, e2023301118 (2021).
124. Bessi, A. et al. Users Polarization on Facebook and Youtube. *PLoS ONE* **11**, e0159641 (2016).
125. Waller, I. & Anderson, A. Quantifying social organization and political polarization in online platforms. *Nature* **600**, 264–268 (2021).
126. Barberá, P. et al. Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Anal.* **23**, 76–91 (2015).

127. Wojcieszak, M., Casas, A., Yu, X., Nagler, J. & Tucker, J. A. Most users do not follow political elites on Twitter; those who do show overwhelming preferences for ideological congruity. *Sci. Adv.* **8**, eabn9418 (2022).
128. Recuero, R., Soares, F. B. & Gruz, A. Hyperpartisanship, disinformation and political conversations on twitter: the Brazilian presidential election of 2018. *Proc. Int. AAAI Conf. Web Soc. Media* **14**, 569–578 (2020).
129. I. Weber, V. R. K. Garimella, and A. Batayneh, “Secular vs. Islamist polarization in Egypt on Twitter,” in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Niagara Ontario Canada: ACM, <https://doi.org/10.1145/2492517.2492557> (2013).
130. Morales, A. J., Borondo, J., Losada, J. C. & Benito, R. M. Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos Interdiscip. J. Nonlinear Sci.* **25**, 033114 (2015).
131. Baqir, A., Galeazzi, A., Drocco, A. & Zollo, F. “Social media polarization reflects shifting political alliances in Pakistan,” 2023. Preprint at <http://arxiv.org/abs/2309.08075> (2024).
132. Takikawa, H. & Nagayoshi, K. Political polarization in social media: Analysis of the ‘Twitter political field’ in Japan. in *2017 IEEE International Conference on Big Data (Big Data)*, Boston, MA: IEEE, 3143–3150 <https://doi.org/10.1109/BigData.2017.8258291> (2017).
133. Nguyen, C. T. et al. Echo chambers and epistemic bubbles. *Episteme* **17**, 141–161 (2020).
134. Terren, L. et al. Open University of Catalonia, R. Borge-Bravo, and Open University of Catalonia, “Echo Chambers on Social Media: A Systematic Review of the literature. *Rev. Commun. Res.* **9**, 99–118 (2021).
135. Sunstein, C. R. The Law of Group Polarization. *J. Polit. Philos.* **10**, 175–195 (2002).
136. Eady, G., Nagler, J., Guess, A., Zilinsky, J. & Tucker, J. A. How many people live in political bubbles on social media? evidence from linked survey and twitter data. *SAGE Open* **9**, 215824401983270 (2019).
137. Di Bona, G. et al., “Sampled social media data risk biased and inconsistent estimates of online social phenomena,” Preprint at SocArXiv. https://doi.org/10.31235/osf.io/37jz9_v1 (2026).
138. Williams, H. T. P., McMurray, J. R., Kurz, T. & Hugo Lambert, F. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Glob. Environ. Change* **32**, 126–138 (2015).
139. Gruz, A. & Roy, J. Investigating political polarization on twitter: a Canadian Perspective: investigating political polarization on twitter. *Policy Internet* **6**, 28–45 (2014).
140. Mekacher, A., Falkenberg, M. & Baronchelli, A. The systemic impact of deplatforming on social media. *PNAS Nexus* **2**, pgad346 (2023).
141. Oswald, L., Schulz, W. S. & Lorenz-Spreen, P. Disentangling participation in online political discussions with a collective field experiment. *Sci. Adv.* **11**, eady8022 (2025).
142. Niezink, N. M. D. “Offline Consequences of Echo Chambers,” in *Proceedings of the Fourth International Workshop on Social Sensing*, Montreal QC Canada: ACM, 5–5 <https://doi.org/10.1145/3313294.3313381> (2019).
143. Guess, A., Nyhan, B., Lyons, B. & Reifler, J. “Avoiding the echo chamber about echo chambers. *Kn. Found.* **2**, 1–25 (2018).
144. Bail, C. et al. *Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing*. (Princeton, NJ: Princeton University Press, 2022).
145. Klein, M., Gruz, A. & Lannigan, J. Using deliberation-centric social network analysis to measure balkanization. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.2914554> (2017).
146. Argyle, L. P. et al. Leveraging AI for democratic discourse: chat interventions can improve online political conversations at scale. *Proc. Natl. Acad. Sci.* **120**, e2311627120 (2023).
147. González-Bailón, S. & Lelkes, Y. Do social media undermine social cohesion? a critical review. *Soc. Issues Policy Rev.* **17**, 155–180 (2023).
148. Tapsell, R. et al. Social media and elections in southeast asia: the emergence of subversive, underground campaigning. *Asian Stud. Rev.* **45**, 117–134 (2021).
149. Oh, C. & Kumar, S. How Trump won: the role of social media sentiment in political elections. *PACIS 2017 Proceedings*, 48, (2017).
150. Mosleh, M., Yang, Q., Zaman, T., Pennycook, G. & Rand, D. G. Differences in misinformation sharing can lead to politically asymmetric sanctions. *Nature* **634**, 609–616 (2024).
151. Guess, A. M. & Munger, K. Digital literacy and online political behavior. *Polit. Sci. Res. Methods* **11**, 110–128 (2023).
152. S. Payne, “Twitter’s bubbles are a blight on British politics,” *The Financial Times*, <https://www.ft.com/content/66df41ad-5835-4f8a-b9f8-7e819722d0c7> (2025).
153. Allcott, H., Braghieri, L., Eichmeyer, S. & Gentzkow, M. The welfare effects of social media. *Am. Econ. Rev.* **110**, 629–676 (2020).
154. Asimovic, N., Nagler, J., Bonneau, R. & Tucker, J. A. Testing the effects of Facebook usage in an ethnically polarized setting. *Proc. Natl. Acad. Sci.* **118**, e2022819118 (2021).
155. Asimovic, N., Nagler, J. & Tucker, J. A. Replicating the effects of Facebook deactivation in an ethnically polarized setting. *Res. Polit.* **10**, 20531680231205157 (2023).
156. Chauchard, S. & Garimella, K. What circulates on partisan WhatsApp in India? insights from an unusual dataset. *J. Quant. Descr. Digit. Media*, <https://doi.org/10.51685/jqd.2022.006> (2022).
157. Shao, C. et al. The spread of low-credibility content by social bots. *Nat. Commun.* **9**, 4787 (2018).
158. Gruz, A., Mai, P. & Soares, F. B. How coordinated link sharing behavior and partisans’ narrative framing fan the spread of COVID-19 misinformation and conspiracy theories. *Soc. Netw. Anal. Min.* **12**, 118 (2022).
159. Yang, K. et al. Arming the public with artificial intelligence to counter social bots. *Hum. Behav. Emerg. Technol.* **1**, 48–61 (2019).
160. J. E. Solsman, “YouTube’s AI is the puppet master over most of what you watch,” *CNET*, <https://www.cnet.com/tech/services-and-software/youtube-cs-2018-neal-mohan/> (2018).
161. Milano, S., Taddeo, M. & Floridi, L. Recommender systems and their ethical challenges. *AI Soc* **35**, 957–967 (2020).
162. Hagen, L., Neely, S., Keller, T. E., Scharf, R. & Vasquez, F. E. Rise of the machines? examining the influence of social bots on a political discussion network. *Soc. Sci. Comput. Rev.* **40**, 264–287 (2022).
163. Stella, M., Ferrara, E. & De Domenico, M. Bots increase exposure to negative and inflammatory content in online social systems. *Proc. Natl. Acad. Sci.* **115**, 12435–12440 (2018).
164. Rettenberger, L., Reischl, M. & Schutera, M. Assessing political bias in large language models. *J. Comput. Soc. Sci.* <https://doi.org/10.1007/s42001-025-00376-w> (2025).
165. Humphreys, D. “AI’s epistemic harm: reinforcement learning, collective bias, and the new AI Culture war,” *Philos. Technol.* <https://doi.org/10.1007/s13347-025-00928-y> (2025).
166. Schroeder, D. T. et al. How malicious AI swarms can threaten democracy. *Science* **391**, 354–357 (2026).
167. Simon F. M. & Altay, S. “Don’t Panic (Yet): assessing the evidence and discourse around generative AI and elections,” KNIGHT FIRST AMENDMENT INSTITUTE, (2025).
168. Hackenburg, K. et al. The levers of political persuasion with conversational AI. *Science* **390**, eaea3884 (2025).
169. Zhou, X., Sharma, A., Zhang, A. X. & Althoff, T. Correcting misinformation on social media with a large language model. *arXiv*

- preprint arXiv:2403.11169, <https://doi.org/10.48550/arXiv.2403.11169> (2024).
170. Costello, T. H., Pennycook, G. & Rand, D. G. Durably reducing conspiracy beliefs through dialogues with AI. *Science* **385**, eadq1814 (2024).
 171. Tessler, M. H. et al. AI can help humans find common ground in democratic deliberation. *Science* **386**, eadq2852 (2024).
 172. Nogara, G. et al. *Toxic Bias: Perspective API Misreads German as More Toxic* <https://doi.org/10.48550/ARXIV.2312.12651> (2023).
 173. Baronchelli, A. Shaping new norms for AI. *Philos. Trans. R. Soc. B Biol. Sci.* **379**, 20230028 (2024).
 174. Burton, J. Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence. *Technol. Soc.* **75**, 102262 (2023).
 175. Persily, N. & Tucker, J. A. Eds., *Social Media and Democracy: The State of the Field, Prospects for Reform*, 1st ed. Cambridge University Press, <https://doi.org/10.1017/9781108890960> (2020).
 176. Fukuyama F. & Grotto A. Comparative Media Regulation in the United States and Europe. in *Social Media and Democracy: The State of the Field, Prospects for Reform*, N. Persily and J. A. E. Tucker, Eds., in *SSRC Anxieties of Democracy*, (Cambridge University Press, 2020).
 177. M. Singh, "India court rejects Twitter's lawsuit against gov't challenging block orders," *TechCrunch*, 2023 <https://techcrunch.com/2023/06/29/court-dismisses-twitter-plea-against-india-government/?guccounter=1> (2024).
 178. Mosleh, M., Martel, C., Eckles, D. & Rand, D. G. Shared partisanship dramatically increases social tie formation in a Twitter field experiment. *Proc. Natl. Acad. Sci.* **118**, e2022761118 (2021).
 179. Roberts, M. et al. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton: Princeton University Press, <https://doi.org/10.23943/9781400890057> (2018).
 180. Farzam, A., Moradi, P., Mohammadi, S., Padar, Z. & Siegel, A. A. Opinion manipulation on Farsi Twitter. *Sci. Rep.* **13**, 333 (2023).
 181. Jones, M. O. Digital authoritarianism in the Middle East: deception, disinformation and social media. London: Hurst & Company, (2022).
 182. Pohle, J. & Thiel, T. "Digital sovereignty," *Internet Policy Rev.* <https://doi.org/10.14763/2020.4.1532> (2020).
 183. Mügge, D. EU AI sovereignty: for whom, to what end, and to whose benefit? *J. Eur. Public Policy* **31**, 2200–2225 (2024).
 184. Wojcieszak, M. & Warner, B. R. Can interparty contact reduce affective polarization? a systematic test of different forms of intergroup contact. *Polit. Commun.* **37**, 789–811 (2020).
 185. Levendusky, M. S. Americans, not partisans: can priming American national identity reduce affective polarization? *J. Polit.* **80**, 59–70 (2018).
 186. Baliotti, S., Getoor, L., Goldstein, D. G. & Watts, D. J. Reducing opinion polarization: effects of exposure to similar people with differing political views. *Proc. Natl. Acad. Sci.* **118**, e2112552118 (2021).
 187. Santoro, E. & Broockman, D. E. The promise and pitfalls of cross-partisan conversations for reducing affective polarization: evidence from randomized experiments. *Sci. Adv.* **8**, eabn5515 (2022).
 188. Harrop, I., Roozenbeek, J., Madsen, J. & van der Linden, S. Inoculation can reduce the perceived reliability of polarizing social media content. *Int. J. Commun.*, **17**, 5291–5315 (2023).
 189. Garibay, I., Mantzaris, A. V., Rajabi, A. & Taylor, C. E. Polarization in social media assists influencers to become more influential: analysis and two inoculation strategies. *Sci. Rep.* **9**, 18592 (2019).
 190. Bail, C. A. et al. Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci.* **115**, 9216–9221 (2018).
 191. Wojcieszak, M. & Garrett, R. K. Social identity, selective exposure, and affective polarization: how priming national identity shapes attitudes toward immigrants via news selection. *Hum. Commun. Res.* **44**, 247–273 (2018).
 192. Gomez, C. J., Herman, A. C. & Parigi, P. Leading countries in global science increasingly receive more citations than other countries doing similar research. *Nat. Hum. Behav.* **6**, 919–929 (2022).
 193. Kahalon, R., Klein, V., Ksenofontov, I., Ullrich, J. & Wright, S. C. Mentioning the sample's country in the article's title leads to bias in research evaluation. *Soc. Psychol. Personal. Sci.* **13**, 352–361 (2022).
 194. Carpentras, D., Lueders, A. & Quayle, M. Response Item Network (ResIN): a network-based approach to explore attitude systems. *Humanit. Soc. Sci. Commun.* **11**, 589 (2024).
 195. M. Quayle, "Social identity networks: People holding attitudes are a collective social identity information system and bipartite networks are a useful way to represent them," *Eur. Rev. Soc. Psychol.* <https://doi.org/10.1080/10463283.2025.2514433> (2025).
 196. Napoli, P. et al. *Social Media and the Public Interest: Media Regulation in the Disinformation Age*. Columbia University Press, <https://doi.org/10.7312/napo18454> (2019).
 197. Merriam-Webster, "Polarization," *Merriam-Webster.com Dictionary*. <https://www.merriam-webster.com/dictionary/polarization> (2026).
 198. Uscinski, J. E. et al. American politics in two dimensions: partisan and ideological identities versus anti-establishment orientations. *Am. J. Polit. Sci.* **65**, 877–895 (2021).
 199. Petersen, M. B., Osmundsen, M. & Arceneaux, K. The 'need for chaos' and motivations to share hostile political rumors. *Am. Polit. Sci. Rev.* **117**, 1486–1505 (2023).
 200. Tucker, J. et al. Social media, political polarization, and political disinformation: a review of the scientific literature. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.3144139> (2018).
 201. LeBas, A. Polarization as craft: party formation and state violence in Zimbabwe. *Comp. Polit.* **38**, 419 (2006).
 202. Pan, J. & Siegel, A. A. How Saudi crackdowns fail to silence online dissent. *Am. Polit. Sci. Rev.* **114**, 109–125 (2020).
 203. LeBas, A. & Gray, K. Authoritarian experience and electoral success: the fate of authoritarian diasporans in Kenya. *Democratization* **28**, 502–520 (2021).
 204. Bowles, J., Larreguy, H. & Liu, S. How weakly institutionalized parties monitor brokers in developing democracies: evidence from postconflict Liberia. *Am. J. Polit. Sci.* **64**, 952–967 (2020).
 205. Arriola, L. R., Devaro, J. & Meng, A. Democratic subversion: elite cooptation and opposition fragmentation. *Am. Polit. Sci. Rev.* **115**, 1358–1372 (2021).
 206. King, G., Pan, J. & Roberts, M. E. How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *Am. Polit. Sci. Rev.* **111**, 484–501 (2017).
 207. Esberg, J. & Siegel, A. A. How exile shapes online opposition: evidence from Venezuela. *Am. Polit. Sci. Rev.* **117**, 1361–1378 (2023).
 208. Tourangeau, R., Edwards, B., Johnson, T. P., Wolter, K. M., & Bates, N. Eds., *Hard-to-Survey Populations*, 1st ed. Cambridge University Press, <https://doi.org/10.1017/CBO9781139381635>. (2014).
 209. Horiuchi, Y., Markovich, Z. & Yamamoto, T. Does conjoint analysis mitigate social desirability bias? *Polit. Anal.* **30**, 535–549 (2022).
 210. Maleki, A. et al. "How do leading methods mislead?: Measuring public opinions in authoritarian contexts," <https://wc2021.ipsa.org/wc/paper/how-do-leading-methods-mislead-measuring-unmeasurable-public-opinions-authoritarian> (2021).
 211. Lukito, J. et al. "The State of Digital Media Data Research, 2023," <https://doi.org/10.26153/TSW/46177> (2023).

212. Narayanan, A. & Zevenbergen, B. No encore for encore? ethical questions for web-based censorship measurement. *SSRN Election. J.* <https://doi.org/10.2139/ssrn.2665148> (2015).
213. Haggard, S. & Kaufman, R. *Backsliding: Democratic Regress in the Contemporary World*, 1st ed. Cambridge University Press, <https://doi.org/10.1017/9781108957809> (2021).
214. Bartolini, S. and Mair, P. *Identity, competition and electoral availability: the stabilisation of European electorates 1885–1985*, 1. publ. Cambridge: Cambridge University Press, (1990).
215. Mair, P. et al. Cleavages. in *Handbook of Party Politics*, SAGE Publications Ltd, 371 (2006).
216. Alba, R. & Foner, N. Immigration and the geography of polarization. *City Community* **16**, 239–243 (2017).
217. Benson, J. et al. Democracy and the epistemic problems of political polarization. *Am. Polit. Sci. Rev.* <https://doi.org/10.1017/S0003055423001089>. (2023).
218. Enyedi, Z. et al. The role of agency in cleavage formation. *Eur. J. Polit. Res.* **44**, 697–720 (2005).
219. Vaishnav, M. et al. “India,” US AID, Contract No. GS-10F-0033M / Order No. 7200AA18M00016, Tasking N068, https://pdf.usaid.gov/pdf_docs/PA0215TX.pdf (2024).
220. Somer, M. et al. Turkey: the slippery slope from reformist to revolutionary polarization and democratic breakdown. *Ann. Am. Acad. Pol. Soc. Sci.* **681**, 42–61 (2019).
221. Basedau, M., Erdmann, G., Lay, J. & Stroh, A. Ethnicity and party preference in sub-Saharan Africa. *Democratization* **18**, 462–489 (2011).
222. Slater, D. & Arugay, A. A. Polarizing figures: executive power and institutional conflict in asian democracies. *Am. Behav. Sci.* **62**, 92–106 (2018).
223. Somer, M. & McCoy, J. Transformations through polarizations and global threats to democracy. *Ann. Am. Acad. Pol. Soc. Sci.* **681**, 8–22 (2019).
224. LeBas, A. & Munemo, N. Elite conflict, compromise, and enduring authoritarianism: polarization in Zimbabwe, 1980–2008. *Ann. Am. Acad. Pol. Soc. Sci.* **681**, 209–226 (2019).
225. Mettler, S. & Lieberman, R. C. *Four Threats: The Recurring Crises of American Democracy*, First edition. New York: St. Martin’s Press, (2020).
226. Abramowitz, A. & McCoy, J. United States: racial resentment, negative partisanship, and polarization in Trump’s America. *Ann. Am. Acad. Pol. Soc. Sci.* **681**, 137–156 (2019).

Author contributions

M.F., M.C., A.G., F.Z., A.B. and W.Q. coordinated the collaboration and collated author contributions. M.F., M.C., A.G., C.A.B., R.M.B., A.Br., A.Gr., D.L., J.K.L., J.M., K.N., D.G.R., A.S., A.A.S., S.v.d.L., O.V., I.W., M.W.,

F.Z., A.Ba., W.Q. all contributed to the writing of the manuscript and approved the final version. A.B. and W.Q. supervised the project.

Funding

Max Falkenberg acknowledges funding from the European Union through the MSCA Actions Postdoctoral Fellowship “POLARCLIMATE”, grant number 101148808.

Competing interests

The Authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-76110-5>.

Correspondence and requests for materials should be addressed to Andrea Baronchelli or Walter Quattrociocchi.

Peer review information *Nature Communications* thanks Naijia Liu, Anna Petherick, Markus Wagner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2026