Article

Political Studies Review
2025, Vol. 23(3) 785–800
© The Author(s) 2024

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14789299241304601
journals.sagepub.com/home/psrev

Can Deliberative Democracy Provide Remedies for Affective Polarisation?

Maija Setälä[1] and Ian O'Flynn[2]

Abstract
This article revisits the problems of electoral democracy, especially affective polarisation, from the perspective of the theory of deliberative democracy. Some authors representing 'realist' views of democracy have concluded that empirical findings regarding affective polarisation challenge not just the idea of representative democracy as responsive government but also the prospects for deliberative democracy. We point out certain problems and limitations in this conclusion and discuss how theories of deliberative democracy might actually help find remedies for affective polarisation. We apply a recent distinction by Hartman et al. as an intermediate-level conceptualisation that helps translate the theory of deliberative democracy into deliberative practices. To illustrate the potential of deliberative practices, we analyse how forums for citizen deliberation such as deliberative mini-publics could help counteract affective polarisation. We reflect on the strengths and weaknesses of deliberative forums from a systemic perspective.

Keywords
deliberative democracy, political behaviour, affective polarisation, deliberative mini-publics

Accepted: 1 November 2024

Introduction

Different 'realist' perspectives on representative democracy have long questioned the degree to which ordinary people are sufficiently informed or reflective to rule themselves (e.g. Downs, 1957). In this same vein, contemporary studies in political psychology show that political preferences and behaviour are, to a large extent, based on affective identifications rather than informed judgements (Achen and Bartels, 2016). In particular, the fast-growing body of empirical research on affective polarisation draws our attention to the tendency, rooted in group identification, of individuals to have strong positive feelings towards their own group combined with negative feelings or feelings of animosity

[1]University of Turku, Turku, Finland
[2]Newcastle University, Newcastle upon Tyne, UK

Corresponding author:
Maija Setälä, Department of Philosophy, Contemporary History and Political Science, University of Turku, Turku 20014, Finland.
Email: maiset@utu.fi

towards the other group (Iyengar et al., 2019). In electoral democracies, affective polarisation based especially on partisan identification has been associated with dysfunctions of representative democracy and democratic backsliding (e.g. Druckman et al., 2023; Wagner, 2024).

On the face of it, affective polarisation seems to apply with particular force to deliberative democracy. As traditionally understood, deliberative democracy is a normative approach to democracy that requires citizens to give reasons for their views and listen seriously to what others have to say in response (e.g. Cohen, 1989: 21; Gutmann and Thompson, 1996). As such, the decisions or conclusions they arrive at should ideally turn not on the force of numbers but on the force of the better argument. However, if, for example, supporters of one party regard the supporters of another in negative terms, if they are effectively closed to what those others have to say, it is on the face of it hard to see how deliberative democracy might ever get off the ground.

The discrepancy between the normative aspirations of deliberative democracy and the realities of politics in representative democracies has long been recognised, including by theorists of deliberative democrats themselves. Already in his early work, Habermas (1989 [1962]) expressed concerns about the role of commercial mass media and its distorting effects on political communication. In the present era, these concerns have been exacerbated by hybrid media systems, featuring alternative media and social media, and their widely discussed role in fuelling affective polarisation and impeding deliberation (Chadwick, 2017; Garrett et al., 2019). When, in the same context, we also consider the rise of populism and extreme right-wing parties, it is perhaps easy to see why many behavioural political scientists are of the view that deliberative democracy is dangerously out of touch with the realities of political behaviour (e.g. Achen and Bartels, 2016: 301; Richey, 2012).

In this article, we therefore address two interlinked questions. First, what implications does affective polarisation have for deliberative democracy? And second, how might deliberative democracy provide remedies for affective polarisation? Despite the sheer amount of empirical research that is now being conducted on affective polarisation, neither question has received sufficient attention in the literature (for an exception, see, for example, Bächtiger and Dryzek, 2024). Accordingly, in what follows, we offer a nuanced account of the ways in which affective polarisation poses a challenge to deliberative democracy and its normative goals. We then show how deliberative democracy can guide our thinking about how democratic systems might be reformed so that some of the negative consequences of affective polarisation can be mitigated or avoided.

To make this latter case, we follow Hartman et al. (2022) who describe interventions to alleviate affective polarisation at three different levels: institutions, relations and thinking. Importantly, these distinctions provide intermediate-level concepts that can be used to link (abstract) deliberative theory and (actual) deliberative practice. Of course, deliberative practices can take different forms in democratic systems. We will focus on how organised citizen deliberation and deliberative mini-publics (Curato et al., 2023), in particular, can be used to counter affective polarisation. This focus should not be read as ignoring the importance other possible reforms and interventions that may also counteract affective polarisation, for example, measures to address economic and political inequalities, the redesign of electoral systems, or media regulation (Bächtiger and Dryzek, 2024: 69–78; Gidron et al., 2020). On the contrary, to be in any way effective in countering affective polarisation, organised citizen deliberation must be approached from a broader

systemic perspective. This imperative stands as an important guiding theme or thread throughout this article.

Affective Polarisation in Electoral Democracies and Its Implications

Different 'realist' perspectives have identified somewhat different problems of electoral democracy. Downs (1957) famously argued that, in mass elections, the costs for any individual voter of acquiring and processing relevant information are likely to be higher than the potential instrumental benefits of voting. Therefore, instrumentally rational voters tend to rely on various cues and shortcuts, especially partisan positions, when forming their opinions and making political choices (Lupia and McCubbins, 1998). More recently, studies have shown the prevalence of motivated reasoning in political life (e.g. Richey, 2012). 'Motivated reasoning' refers to variety of cognitive and affective mechanisms which lead individuals to aim at the conclusions they want to arrive at (Kunda, 1990), by using their capacity for reasoning to defend evidence supporting their beliefs while discounting or dismissing contradictory evidence (Taber and Lodge, 2006). Importantly, motivated reasoning is often rooted in group identification. In the context of current representative democracies, partisan group identification seems to explain people's political perceptions and views – rather than the other way around (Achen and Bartels, 2016; Slothuus and de Vreese, 2010).

While partisan motivated reasoning poses obvious challenges to certain conceptions of democracy, including deliberative democracy, our focus in this article is on *affective polarisation* (e.g. Druckman et al., 2023; Iyengar et al., 2019). Affective polarisation is defined as a combination of positive feelings towards ingroup and negative feelings or feelings of animosity towards outgroups. In representative systems, affective polarisation can be based on partisan identification, but also, it should be said, on other social cleavages such as class, ethnicity, religion, (non-party-based) ideology and moral convictions or positions on crucial identity issues (Hobolt et al., 2021). Affective polarisation may be related to ideological polarisation, but this is not necessarily always the case (Iyengar et al., 2019).

Focusing specifically on partisan group identification, studies suggest that affective polarisation across party lines is heightened during electoral campaigns (Hernandez et al., 2021), confirming the effects of institutional context in this respect. Unsurprisingly, negative campaigning in elections has been found to increase affective polarisation (Martin and Nai, 2024), highlighting the discursive aspects of the phenomenon. Evidence from the United States suggests that affective polarisation among voters is increasing and there are deepening divisions across party lines (e.g. Iyengar et al., 2019; Iyengar and Westwood, 2015). Studies on affective polarisation in European systems with proportional representation suggest that coalition governments can lessen partisan animosity between supporters of coalition partners (Wagner, 2024). However, these studies also highlight the prevalence of affective polarisation between different partisan groups or 'affective blocs' (Reiljan, 2020; Wagner, 2024).

Affective polarisation based on partisan identifications is by no means a new phenomenon in representative democratic systems. Yet, increased scholarly attention seems to reflect its dramatic rise in recent decades (see, for example, Wagner, 2024). There is much scholarly debate on the reasons underlying this recent rise, ranging from growing socio-economic inequalities to new information and communication technologies and their

effects on media systems and political campaigning. There are also explanations based on changes in patterns of distribution of political power, for example, the non-democratic tendencies in increasingly technocratic policymaking (White and Neblo, 2021).

For some, the role of affective identifications in political reasoning and the prevalence of affective polarisation should change our normative expectations of democracy. For example, in *Democracy for Realists*, Achen and Bartels (2016) conclude that the role of affective identification in voters' reasoning is such that ideas of democracy as popular self-government – or government which is responsive to people's preferences – should be abandoned. As part of this, Achen and Bartels interpret studies of political behaviour so that they prove false, or demonstrate the futility of, those conceptions of democracy that entail higher than minimal expectations of electoral democracy, including theories of deliberative democracy.

While the findings of studies in political behaviour are obviously relevant to democratic theory, the conclusion by Achen and Bartels' seems to be unsatisfactory because affective polarisation may undermine the stability of electoral democracy itself. Empirical studies suggest that affective polarisation may create room for parties and candidates who are not committed to democratic norms (Graham and Svolik, 2020; Iyengar et al., 2019). Moreover, comparative studies find that strong affective (non-ideological) polarisation is associated with democratic backsliding globally (Orhan, 2022). Admittedly, there is also evidence supporting different interpretations of the effects of affective polarisation on democratic norms. For example, affective polarisation against radical right parties can be associated with support for democratic norms (Meléndez and Kaltwasser, 2021), and some experimental studies find no association between affective polarisation and attitudes that undermine democratic norms (Broockman et al., 2023). Further empirical research is thus needed to specify the effects of affective polarisation on democratic norms and the stability of electoral democracy. However, our immediate concern, to which we will now turn, is with the implications of affective polarisation for the theory of deliberative democracy and the normative aspirations with which it is bound up.

Implications of Studies on Affective Polarisation for the Theory of Deliberative Democracy

To begin with, it is noteworthy that the normative standards of deliberative democracy are obviously different from those interpretations of democracy which emphasise fair electoral competition and regular government turnover(e.g. Schumpeter, 1962 [1942]). For example, Bächtiger et al. (2018: 2) define democratic deliberation as a practice, where 'people come together, on the basis of equal status and mutual respect, to discuss the political issues they face and, on the basis of those discussions, decide on the policies that will then affect their lives'. Equal, open and inclusive deliberation is thought necessary for collective will-formation which, from a deliberative perspective, is key to both the democratic legitimacy of policies and citizens' commitment to collective goals (see, e.g. Warren, 2020). Rather than the idea of responsive government criticised by Achen and Bartels, deliberative democracy emphasises the possibility of popular self-government based on deliberative processes of collective will-formation (e.g. Dryzek and List, 2003).

That said, empirical findings regarding affective identification and polarisation clearly bear on these key normative ideas of deliberative democracy. First, empirical findings suggest that people tend to form their political judgements and evaluations based on affective identifications rather than on the quality or merits of the arguments that they

hear. Reasoning guided by partisan and other identifications seems to be directly at odds with the ideas of (equal, inclusive, open-minded) reasoning, listening and reflection presupposed in theories of deliberative democracy. There are some theoretical attempts to develop more 'realist' versions of deliberative democracy which would recognise affective identifications as an element of public reasoning (cf. Lafont, 2019). These attempts seem to be compatible with the view that people need identity-based cues and shortcuts to navigate and make sense of the complexities of politics (Leeper and Slothuus, 2014). However, studies on political behaviour also draw our attention to the risks involved in political reasoning guided by affective identification, such as mass-scale errors in political judgement and extreme (ideological and affective) polarisation. These negative consequences highlight the importance of various corrective mechanisms such as democratic deliberation in representative systems, but also obstacles to them (more on which below).

Second, and even more crucially, affective polarisation refers not just to positive feelings towards an ingroup, but also to negative feelings towards outgroups – or partisan animosity. Partisan animosity is often related to negative perceptions regarding the character, intentions or reliability of those representing other parties or outgroups more generally. Partisan animosity has impacts not only on political attitudes, but also in terms of social relations. These 'downstream' impacts of partisan animosity bear, for example, on individuals' preferences and choices regarding where they want to live as well as with whom to work and socialise. There is evidence that partisan animosity can lead to the erosion of social and political trust, and hostility between different groups (e.g. Iyengar and Westwood, 2015).

From the perspective of deliberative democracy, these 'downstream' impacts seem especially problematic. Along with other forms of social sorting, partisan animosity is likely to decrease people's willingness to engage in deliberation with the other side (Strickler, 2018). Under conditions of affective polarisation and partisan animosity, in particular, people are not just dismissive of political arguments and views put forward by 'outgroups', but also unwilling to engage or communicate with them in the first place. Partisan animosity, therefore, limits the prospects for deliberative exchanges of reasons across social and political cleavages and hence undermines the prospects for deliberative democracy more generally (cf. Benson, 2024).

There are, however, other ways of understanding the implications of affective polarisation from the perspective of deliberative democracy understood as a more complex, macro-political system. According to Habermas' (1996) 'two-track' model, there is (or normatively ought to be) a division of deliberative labour between a public sphere centred on the institutions of civil society, on the one hand, and a political sphere centred on parliamentary decision-making procedures on the other – or, in Nancy Fraser's (1992) terms, between 'weak publics' and 'strong publics'. On this view, weak and strong publics are meant to engage in a productive dialogue. In the first instance, weak publics are responsible for identifying and articulating policy problems and for generating public opinion on these matters, while strong publics are responsible for filtering public opinion through formal parliamentary procedures and for arriving at collectively binding decisions.

From such a macro-systemic perspective, affective polarisation may even be perceived in a positive light since it has been found to have politically mobilising effects (Harteveld and Wagner, 2023). Affective polarisation could energise or revitalise the public sphere, especially if it were to foster or facilitate the articulation of marginalised and disempowered viewpoints or (as mentioned above) to defend democratic norms. On balance, however, mobilisation based on partisan animosity is likely to be undesirable from the

perspective of deliberative democracy; for example, anger may motivate and mobilise, but it may also lead partisans to think only in terms of defeating the other side, rather than in terms of mutually justifiable political change (Druckman et al., 2023). Moreover, partisan animosity undermines the willingness to compromise as well as reduces the capacity for communication, listening and perspective-taking across cleavages. Although affective polarisation may be inevitable and sometimes even valuable for democracy, from the normative perspective of deliberative democracy, its potentially positive effects can only take us so far.

Affective polarisation, and partisan animosity in particular, is therefore largely problematic when it comes to deliberative democracy understood as a macro-political system based on open and inclusive reasoning across social and political cleavages. This concern also applies to more 'realistic' views of the cognitive demands of deliberative democracy. For example, according to Warren (2020: 86), 'it is important in a democracy that individuals be able to prioritise their attention, so they can focus on the issues that are most important to them'. On this view, a great deal of deliberative labour must occur within representative institutions broadly understood. Individuals must be able to have trust in representatives and be able to hold them accountable when there is reason for distrust. However, trust in representatives poses problems of its own.

Trust in elected and other representatives should be based on citizens' informed and considered judgements on whether representatives' values and interests are congruent with their own, as well as on whether their actions match their words (MacKenzie and Warren, 2012; cf. Warren and Gastil, 2015). Likewise, meaningful deliberative accountability presupposes that citizens are able to make good judgements about representatives' actions and the justifications that they give for those actions. However, as Druckman et al. (2023) point out, as a consequence of affective polarisation, trust may be politicised to the extent that people only trust their 'own' party and interpret political facts and preferences accordingly. Consequently, citizens may assess their representatives' performance on the basis of affective identifications or misplaced attributions of credit or blame (Achen and Bartels, 2016; cf. Warren, 2020: 86). Therefore, because of the risks of 'blind' trust and accountability based on affective identifications, even these more 'realistic' views on the role of public reasoning in representative systems are not immune to the problems caused by affective polarisation.

To sum up the discussion thus far, empirical findings suggest that affective identifications and affective polarisation, especially partisan animosity, undermine the prospects for democratic deliberation. Because these studies conclude that that people are 'not psychologically wired to listen to the other side, to be open to persuasion, or to offer sincere reasons' (Chambers, 2018: 38), the key normative goals of deliberative democracy seem largely unachievable. However, it is also notable that these empirical findings are made in the context of electoral representative systems. While the tendencies towards partisan motivated reasoning and affective polarisation may be 'ingrained' (to an appreciable extent) in human nature, the current institutional structures of electoral democracy seem to aggravate these tendencies.

Deliberative democratic theorists have long argued that institutions matter to the likelihood of deliberative processes, and that different institutional conditions are more or less favourable for more deliberative forms of reasoning (Cohen, 1989: 26). The problem is not that people are unable to deliberate; on the contrary, there is evidence of people's capacity to deliberate in particular contexts. For example, Landemore (2020: 45–46) has pointed out that, while Achen and Bartels may be right that electoral democracies are not

deliberative in character, this conclusion should not be regarded as a reason to abandon the normative ideas of deliberative democracy. Because deliberative democracy is, after all, a normative project, this conclusion should instead be regarded as a reason to look for other ways of organising democratic systems to avoid these problems. If electoral institutions undermine prospects of democratic deliberation by giving rise to affective polarisation, we should look for other institutional conditions that are more favourable for democratic deliberation.

Reforming current democratic systems – again, understood here as a macro-political project – requires a thorough understanding of the institutional conditions that hinder or support more deliberative forms of reasoning. In this spirit, Chambers (2023: 3) argues that 'It is important to know why citizens are polarized and why they are victims of disinformation. But this surely is not as important a question for democracy as how to get past polarisation and disinformation'. However, in our view, these two questions – why people are polarised and how we can get past it – are essentially interlinked. In particular, it is necessary to understand the mechanisms behind affective identifications and polarisation in electoral democracies in order to identify practices of democratic deliberation that, from a systemic perspective, can help avoid the harmful consequences or work as antidotes to them.

The Possibility of Forums for Citizen Deliberation to Counteract Affective Polarisation: Institutions, Relations and Thinking

Given widespread concerns regarding the current state of representative democracy, deliberative democrats are by no means alone in trying to identify ways to address its problems. Indeed, there are a variety of proposals for institutional designs and interventions aimed at avoiding the pathological consequences of affective polarisation and partisan animosity in particular. Such designs and practices are valuable also from the perspective of deliberative democracy to the degree that they help eradicate obstacles to open and inclusive reasoning and reflection.

For example, in a major review of the literature, Hartman et al. (2022: 1195) note that '[p]olitical and psychological scientists have outlined several frameworks to synthesize the causes of partisan animosity, ranging from personal thoughts to interpersonal relationships and societal institutions'. Partisans tend to hold negative *thoughts* about their political opponents; *relations* are further undermined by sorting along multiple social dimensions; while the structure of public *institutions* plays its part in reinforcing negative feelings about other groups. In response, Hartman et al. (2022) show how many promising interventions for reducing partisan animosity can be categorised under these same three conceptual headings. They point, for example, to a broad range of studies in social psychology showing how positive contact and communication across partisan lines corrects misconceptions and stereotypes (thoughts) and highlights commonalities (relations), and to studies in political science and communication studies showing how electoral reform or media regulation (institutions) might also play their part.

Much of what these authors have to say is consistent with goals of deliberative democracy. Among other things, deliberative theory has long pointed to the importance of carefully curated information, building discussion skills, fostering communication and positive contact across cleavages, reforming governmental structures and bolstering the

public sphere (see, for example, Bächtiger and Dryzek, 2024; Gastil and Knobloch, 2020). In what follows, we will therefore use the institutions, relations, thoughts distinction as an intermediate-level conceptual framework for thinking about how different forms of democratic deliberation might effectively respond to the problems of affective polarisation and partisan animosity. However, as indicated in our introduction, while intermediate-level conceptualisation of this sort can help us to see how normative deliberative theory could be relevant for addressing affective polarisation, to be really impactful they also need to be located within a broader systemic view of democratic practices.

We will therefore apply Hartman et al.'s distinction to evaluate the capacity of forums for organised citizen deliberation to counteract affective polarisation and to explore the constraints and opportunities of deliberative forums from a systemic perspective. We will discuss different forms of organised citizen deliberation, and deliberative mini-publics in particular. A deliberative mini-public is a small group of citizens selected at random from a larger population and tasked with deliberating about an important matter of law or public policy. To this end, participants are given the time and space to learn about the issues – for example, by interacting with experts – and to deliberate in facilitated small groups with others representing different viewpoints (Curato et al., 2023; Setälä and Smith, 2018). Group composition, especially diversity of viewpoints, is crucial for both enhancing critical thinking and avoiding group thinking (Mercier and Landemore, 2012: 253). Participants, under the guidance of trained moderators, both justify their own arguments and respond to those put to them by others.

The effects of such reasoning processes are manifold. Participants not only learn about the issue under discussion, but also develop their capacity for perspective-taking (Grönlund et al., 2017); even in deeply divided societies, they tend to see one another as open to reason and regard joint problem-solving as the norm (Luskin et al., 2014). Granted, there are also other forums for public deliberation besides randomly selected mini-publics, for example, parliamentary and select committees, town hall meetings and the like. But it seems that the combination of random selection and facilitated small-group discussions is a particularly efficient remedy for problems such as motivated reasoning and affective polarisation. Importantly, random selection effectively prevents the possibility that deliberators use partisan and other identifications as shortcuts, thus encouraging careful listening and more complex forms of information processing and reasoning.

Against this backdrop, mini-publics seem to have the potential to address polarising issues such as climate policies. There is already a wealth of empirical evidence showing that mini-publics can foster inclusive deliberation and develop informed and considered judgements regarding policies (cf. Setälä and Smith, 2018). From the perspective of the theory of deliberative democracy, the main objective of practical deliberative designs is to help people make better judgements, to improve the quality of public justifications and reasoning and, consequently, to make better and more legitimate public decisions. Though the link to decision-making at the level of actual public policy can be indirect, or hard to track, there is evidence that deliberative mini-publics can play a pivotal role (e.g. Doyle and Walsh, 2024).

However, the capacity of mini-publics to enhance public reasoning in representative systems *more broadly*, and especially among *the wider public*, remains debatable. Although deliberative mini-publics seem to foster deliberation and act as an antidote to affective polarisation among participants, there are still many open questions about their capacity to deal with affective polarisation more generally – that is, as we have stressed

throughout, as a macro-political or systemic problem (Rountree et al., 2022). After all, mini-publics are usually convened on an ad hoc basis for a relatively short period of time and (most obviously of all) directly engage only a tiny fraction of the population from which participants are drawn.

We will next explore how forums for citizen deliberation and mini-publics, in particular, can be situated within representative systems in order to influence communication, information processing and reasoning in the context of mass publics. We are particularly interested in the ways in which they may potentially counteract affective polarisation across partisan lines and aggravated by electoral competition. Building on Hartman et al.'s (2022) approach, we highlight constraints and opportunities related to different ways of organising deliberative forums. In doing so, we will refer to empirical evidence on the effects of these interventions wherever available, though, admittedly, in some cases, the relevant evidence of the effects is still largely speaking missing.

Institutions

Inspired by these results, Landemore (2020) has proposed that the role of randomly selected bodies should be (dramatically) enhanced in democratic systems, not just as supplements to existing electoral institutions but also as an alternative to them. Following the model of the ancient Athenian democracy (e.g. Ober, 1989), representatives would be selected by lot rather than by voting and would serve for a fixed period of time before being rotated. On the face of it, this radical view promises to undercut the realist critique of democracy (cf. Achen and Bartels, 2016). By replacing elections with random selection as a mechanism for selecting representatives, the characteristics of the public sphere and political discourse would be radically changed. There would be no need for citizens to use partisan cues to make their choices in elections, and hence one of the main drivers of affective polarisation would be dissolved.

While lottocratic systems such as Landemore's 'open democracy' might foster inclusive reasoning and deliberation among those selected, there is a risk that such systems would entail an array of other problems that are actually more serious than those caused by partisan competition in elections. As pointed out earlier, party identification is not just a driver of motivated reasoning and affective polarisation, but is also a form of shortcut that helps people make sense of the complexities of politics. In this respect, the absence of parties as sources of affective identifications and shortcuts in lottocratic systems might even be detrimental to democratic deliberation in the public sphere (e.g. Rummens and Geenens, 2023). From a broader perspective, electoral competition seems to be a key safeguard against abuses of power because it drives political engagement and mobilisation.

Instead of speculating about the consequences of such radical reforms, we focus on potential ways in which randomly selected bodies can supplement electoral representation by counteracting the negative aspects of affective polarisation. Compared to the proposal to replace elections with lottocracy, these proposals entail more modest institutional reforms, or limited interventions, within existing institutional structures. The principal aim of these interventions is to counterbalance electoral representation and to capitalise on the benefits of both elected and randomly selected institutions. For example, both Abizadeh (2021) and Gastil and Wright (2018) defend the idea of randomly selected second chambers in bicameral systems. Such institutional designs could enhance inclusion and 'foster the kind of conscientious deliberation stymied by strategic and electoral

imperatives' (Abizadeh, 2021: 798). Although there are an increasing number of permanent deliberative forums attached to legislative institutions (cf. Macq and Jacquet, 2023), randomly selected second chambers have not been put into practice and therefore their effects on the wider public remain largely speaking speculative.

Another potentially efficacious way of enhancing deliberation in representative politics is 'directly representative democracy' (Neblo et al., 2018). In essence, the proposal here is to establish a new form of congressional town-hall meeting between a member of congress and a randomly selected cross-section of constituents. Such 'ongoing deliberative accountability' encourages elected representatives to deliberate with audiences that are broader than their own core voters. Crucially, the type of interaction envisaged here could help avoid polarisation and enhance the quality of justification among elected representatives, not just because it encourages representatives to consider a broader range of views and interests, but also because it encourages them to do so on an ongoing basis.

The obvious problem with proposals for randomly selected second chambers and directly representative democracy is that they do not *directly* affect modes of communication and reasoning among the public at large. Therefore, their capacity to address the problem of affective polarisation may remain limited, especially in the context of electoral competition and campaigning. Nevertheless, randomly selected second chambers and directly representative democracy could enhance the quality and the inclusiveness of public justifications, create new venues for deliberative interactions between citizens and representatives, and – most importantly of all – mitigate elite polarisation. Therefore, viewed from a more systemic perspective, they could *indirectly* help improve the quality of public discourse overall and in this way mitigate the effects of affective polarisation. To what degree they might do so is at this point an open question; yet, a systemic perspective at least enables us to see some value in initiatives of this sort, taken in the round.

Relations

In their 'America in One Room' study, Fishkin et al. (2021) find empirically that organised deliberation can be an antidote both to opinion and affective polarisation among participants. In this way, organised democratic deliberation seems to work in similar ways to initiatives that have been designed to foster positive contact across social cleavages, for example, 'More in Common UK' and 'Kentucky Rural Urban Exchange'. However, these models can foster communication and positive contact only among the small number of people involved, and their effects may remain negligible from a systemic perspective.

Though proposed some 20 years ago, Ackerman and Fishkin's (2002) 'Deliberation Day' remains therefore an important touchstone. Deliberation Day is intended as a national holiday, held 1 week before major national elections, allowing registered voters to gather together in neighbourhood meeting places to deliberate about the issues at stake in the election. As in mini-publics, participants have opportunities to learn about the issues and the deliberative process is facilitated. In order to encourage participation, voters are financially compensated for their time, on condition that they subsequently turn out to vote. From the perspective of our argument, the key advantage of the idea of the Deliberation Day is that it would be connected to elections, which often heighten affective polarisation. In this way, Deliberation Day is designed to help voters make informed and reflected political choices and not simply to foster communication across cleavages. Deliberation Day would also not be 'only' a randomly selected deliberative mini-public since it would also engage (nearly) the whole voting-age population.

The obvious problem with this idea is, however, that neighbourhood meetings may not be diverse – partly because the neighbourhoods themselves may not be especially socially diverse; in fact, they may be shaped by mechanisms such as partisan sorting. Therefore, the problem remains that of fostering communication across social cleavages. Online deliberation could be one possibility for overcoming physical restraints and facilitating communication across geographical boundaries. In this regard, some of the more intriguing developments relate to the use of artificial intelligence (AI). For example, the Stanford Online Deliberation Platform (Siu and Wakao, 2022) uses an AI facilitator to moderate deliberations, ensuring, among other things, that no participant dominates the deliberations. While AI facilitators may not do a better job than humans (Gelauff et al., 2023), they are certainly much cheaper, which matters if we are ever to utilise processes such as the Deliberation Day to enhance inclusive deliberation among large-scale publics.

Thoughts

There are also ways of using mini-publics as trusted sources of information which help learning and reflection among voters (Gastil and Knobloch, 2020). A case in point is the Citizens' Initiative Review (CIR), which has been designed to help voters make better informed and reflected judgements on ballot initiatives (Gastil and Richards, 2013). In this design, a Citizens' Jury meets to discuss the measure on the ballot and produces a written statement containing a description of the jury, key facts and the most important reasons for and against the measure. This statement is then distributed to voters with the intention of helping them to cast an informed vote. By setting out the key facts as well as the best arguments pro and con, the CIR lowers voters' cognitive costs of acquiring and processing information (Warren and Gastil, 2015: 566–567). Indeed, there already is some empirical evidence to suggest that CIR-type mini-publics can help voters make informed judgements and, moreover, enhance learning and reflection (e.g. Knobloch et al., 2020).

While the CIR seems to be comparable to interventions to counter misinformation, it can also counteract affective polarisation by enhancing cross-cutting communication during political campaigns. Importantly, because it introduces voters to arguments that they might otherwise not encounter, the CIR provides an example of a mini-public that may help engender inclusive reasoning, listening and reflection (Lafont, 2019). For this reason, Gastil and Richards (2013) have described the CIR as a voting aid rather than a shortcut or a cue. Furthermore, some authors (e.g. Drury and Rountree, 2024) have suggested that CIR statements encourage 'vicarious deliberation'; when organised on a polarised issue, the description of the jury process and its outcomes encourages voters to think about a situation where people representing different sides of the conflict can communicate, listen to and evaluate each others' arguments.

In short, reading a CIR statement could counteract affective polarisation by exposing the readers to the arguments and rationales on the other side. Although the evidence is still limited, there are some studies giving support to this view. A study by Már and Gastil (2020) finds that reading a CIR statement can help attenuate the effects of partisan motivated reasoning, and a study by Gastil et al. (2023) suggests that it can spark accuracy motivation. Moreover, reading a CIR statement can also enhance consideration of conflicting viewpoints even in contexts where affective polarisation undermines trust in information provided by media and public institutions (Setälä et al., 2023).

While the CIR was designed to review ballot initiatives, similar procedures could be used to review policy issues submitted to other types of referendum. Moreover, CIR-type interventions could even be used to help voters make informed and balanced judgements on issues discussed in electoral campaigns (cf. Gastil, 2000). Admittedly, the effects observed in conjunction with popular votes may be harder to achieve in elections where there are entire policy platforms (as opposed to individual policy measures) at stake and where partisan sentiments may play a stronger role – and where a polarised media may be more actively engaged. Nevertheless, even modest effects might be important when taken in the round. When properly designed mini-publics could be used to help voters to evaluate factual and arguments in electoral campaigns. This could help voters to understand the views and rationales of those on the other side, and appeals to partisan identifications and partisan animosity might lose some of their force in campaign rhetoric.

Obviously, our review of possible deliberative democratic designs is not comprehensive, but it is intended to illustrate the kinds of issues and concerns that arise when considering such designs as possible remedies for the negative consequences of affective polarisation. Some of the proposals entail a radical overhaul of systems based on electoral representation, whereas others aim at correcting, or at least ameliorating, the negative consequences of existing practices and behaviour in electoral democracies. The question of how the suggested measures actually work is empirical in character, and in some cases, the empirical evidence is currently scarce (and may remain so). Nevertheless, out review shows that the theory of deliberative democracy can be helpful in terms of developing institutional designs and interventions that counteract to the negative aspects of affective polarisation.

Concluding Comments

Affective polarisation is probably an inevitable element in democratic politics. The affective identifications with which it is bound up are arguably essential for making sense of politics, engagement and articulation of interests. Although affective polarisation seems to contribute to the erosion of elite cooperation and political civility more broadly, the evidence on its effects for the support of democratic norms is mixed. What we can say, however, is that affective polarisation impedes the functioning of representative democracy as currently practised. Affective polarisation, and partisan animosity in particular, may lead to distrust among people with different political identifications, and hence to ‘tribal politics’. Affective identifications may also lead to misguided judgements regarding trustworthiness and ‘blind’ accountability.

Affective polarisation also undermines the prospects for democratic deliberation. Partisan animosity is particularly harmful for the democratic public sphere because it prevents communication and mutual justification across social and political cleavages. Deliberative democrats – or anybody concerned about the quality of democracy – should be aware of these risks and should also be interested in ways in which they can be reduced. There should be efforts to identify societal factors that might explain the current increase in affective polarisation, not least because this helps us to identify and design institutions and interventions that counteract its negative consequences.

Deliberative theory may be helpful in this respect – especially when linked, via intermediate-level concepts which link theory to practice – by showing ways to support processes of mutual justification across cleavages and more reflective and inclusive forms of reasoning among citizens. Reforming democratic systems requires finding a balance

between institutional designs that encourage political engagement and designs that encourage inclusive and reflective forms of reasoning. From the perspective of deliberative democracy, there are many possible routes to alleviate affective polarisation at the levels of institutions, relations and thinking. These include broader reforms of political institutions and the media system (e.g. Bächtiger and Dryzek, 2024), measures to foster cross-cutting contact and dialogue (Hendriks et al., 2020), and deliberative journalism reforms (e.g. Carcasson, 2022).

However, it is also important to keep in mind the potential of forums for organised citizen deliberation. The primary goal of deliberative designs such as mini-publics is to improve the quality of reasoning and judgements in public decision-making. If adequately connected to the key democratic processes, and electoral competition in particular, forums for citizen deliberation can also play their part in countering affective polarisation. Empirical studies on CIR processes already provide some promising results showing how small-scale deliberative bodies can have large-scale impacts on voters' thinking. Given the stakes, we have every reason to further explore the potential systemic effects of such deliberative designs.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research has been funded by The Research Council of Finland (project FACTOR, decision number 341373).

ORCID iD

Maija Setälä https://orcid.org/0000-0002-9341-5584
Ian O'Flynn https://orcid.org/0000-0002-4526-2739

References

Abizadeh A (2021) Representation, Bicameralism, Political Equality, and Sortition: Reconstituting the Second Chamber as a Randomly Selected Assembly. *Perspectives on Politics* 19 (3): 791–806.

Achen CH and Bartels LM (2016) *Democracy for Realists*. Princeton, NJ: Princeton University Press.

Ackerman B and Fishkin JS (2002) Deliberation Day. *The Journal of Political Philosophy* 10 (2): 129–152.

Bächtiger A and Dryzek J (2024) *Deliberative Democracy for Diabolical Times*. Cambridge: Cambridge University Press.

Bächtiger A, Dryzek J, Mansbridge J, et al. (2018) Deliberative Democracy: An Introduction. In: Bächtiger A, Dryzek J, Mansbridge J and et al. (eds) *The Oxford Handbook of Deliberative Democracy*. Oxford: Oxford University Press, pp.1–31.

Benson J (2024) Democracy and the Epistemic Problems of Political Polarization. *American Political Science Review* 118: 1719–1732.

Broockman DE, Kalla JL and Westwood SJ (2023) Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not. *American Journal of Political Science* 67: 808–828.

Carcasson M (2022) Taking on Two Crises: Democracy and Journalism. *National Civic Review* 111 (3): 16–26.

Chadwick A (2017) *The Hybrid Media System: Politics and Power*, 2nd ed. Oxford: Oxford University Press.

Chambers S (2018) Human Life Is Group Life: Deliberative Democracy for Realists. *Critical Review* 30 (1–2): 36–48.

Chambers S (2023) Citizenship in Times of Crisis: A Comment on Danielle Allen's Democratic Theory. *Polity* 55 (4): 726482.

Cohen J (1989) Deliberation and Democratic Legitimacy. In: Hamlin A and Pettit P (eds) *The Good Polity*. Oxford: Blackwell, pp.17–34.

Curato N, Farrell D, Geißel B, et al. (2023) *Deliberative Mini-Publics: Core Design Features*. Bristol: Bristol University Press.

Downs A (1957) *An Economic Theory of Democracy*. New York: Harper & Brothers.

Doyle O and Walsh R (2024) Assessing the Influence and Legitimacy of Citizen Deliberation on Abortion: A Reply to Eoin Carolan and Seana Glennon. *International Journal of Constitutional Law* 22: 204–208.

Druckman JN, Green DP and Iyengar S (2023) Does Affective Polarization Contribute to Democratic Backsliding in America? *The Annals of the American Academy of Political and Social Science* 708 (1): 137–163.

Drury SA and Rountree J (2024) The Genre of Deliberative Guidance: Rhetoric and Deliberation in Citizens' Initiative Review Statements. *Western Journal of Communication* 88: 125–146.

Dryzek J and List C (2003) Social Choice Theory and Deliberative Democracy: A Reconciliation. *British Journal of Political Science* 33 (1): 1–28.

Fishkin J, Siu A, Diamond L, et al. (2021) Is Deliberation an Antidote to Extreme Partisan Polarization? Reflections on 'America in One Room'. *American Political Science Review* 115 (4): 1464–1481.

Fraser N (1992) Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. In: Calhoun C (ed.) *Habermas and the Public Sphere*. Cambridge, MA: MIT Press, pp.109–142.

Garrett RK, Long JA and Jeong MS (2019) From Partisan Media to Misperception: Affective Polarization as Mediator. *Journal of Communication* 69 (5): 490–512.

Gastil J (2000) *By Popular Demand: Revitalizing Representative Democracy through Deliberative Elections*. Berkeley, CA: University of California Press.

Gastil J and Knobloch KR (2020) *Hope for Democracy: How Citizens Can Bring Reason Back into Politics*. Oxford: Oxford University Press.

Gastil J and Richards R (2013) Making Direct Democracy More Deliberative through Random Assemblies. *Politics & Society* 41 (2): 253–281.

Gastil J and Wright EO (2018) Legislature by Lot: Envisioning Sortition within a Bicameral System. *Politics & Society* 46 (3): 303–330.

Gastil J, Már Ársælsson K, Knobloch KR, et al. (2023) Deliberative Panels as a Source of Public Knowledge: A Large-Sample Test of the Citizens' Initiative Review. *PLoS ONE* 18 (7): e0288188.

Gelauff L, Nikolenko L, Sakshuwong S, et al. (2023) Achieving Parity with Human Moderators: A Self-Moderating Platform for Online Deliberation. In: Boucher S, Hallin CA and Paulson L (eds) *The Routledge Handbook of Collective Intelligence for Democracy and Governance*. New York: Routledge, pp.202–221.

Gidron N, Adams J and Horne W (2020) *American Affective Polarization in Comparative Perspective*. Cambridge: Cambridge University Press.

Graham MH and Svolik MW (2020) Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States. *American Political Science Review* 114 (2): 392–409.

Grönlund K, Herne K and Setälä M (2017) Empathy in a Citizen Deliberation Experiment. *Scandinavian Political Studies* 40: 457–480.

Gutmann A and Thompson D (1996) *Democracy and Disagreement*. Cambridge, MA: The Belknap Press.

Habermas J (1989 [1962]) *The Structural Transformation of the Public Sphere*. Cambridge: Polity Press.

Habermas J (1996) *Between Facts and Norms*. Cambridge, MA: MIT Press.

Harteveld E and Wagner M (2023) Does Affective Polarisation Increase Turnout? Evidence from Germany, The Netherlands and Spain. *West European Politics* 46 (4): 732–759.

Hartman R, Blakey W, Womick J, et al. (2022) Interventions to Reduce Partisan Animosity. *Nature Human Behaviour* 6: 1194–1205.

Hendriks CM, Ercan S and Boswell J (2020) *Mending Democracy: Democratic Repair in Disconnected Times*. Oxford: Oxford University Press.

Hernandez E, Anduiza E and Rico G (2021) Affective Polarisation and the Salience of Elections. *Electoral Studies* 69: 102203.

Hobolt SB, Leeper TJ and Tilley J (2021) Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum. *British Journal of Political Science* 51 (4): 1476–1493.

Iyengar S and Westwood SJ (2015) Fear and Loathing across Party Lines: New Evidence on Group Polarization. *American Journal of Political Science* 59 (3): 690–707.

Iyengar S, Lelkes Y, Levendusky M, et al. (2019) The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science* 22: 129–146.

Knobloch K, Barthel M and Gastil J (2020) Emanating Effects: The Impact of the Oregon Citizens' Initiative Review on Voters' Political Efficacy. *Political Studies* 68 (2): 426–445.
Kunda Z (1990) The Case for Motivated Reasoning. *Psychological Bulletin* 108 (3): 480–498.
Lafont C (2019) *Democracy Without Shortcuts: A Participatory Conception of Deliberative Democracy*. Oxford: Oxford University Press.
Landemore H (2020) *Open Democracy: Reinventing Popular Rule for the Twenty-First Century*. Princeton, NJ: Princeton University Press.
Leeper TJ and Slothuus R (2014) Political Parties, Motivated Reasoning, and Public Opinion Formation. *Advances in Political Psychology* 35: 129–156.
Lupia A and McCubbins MD (1998) *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge: Cambridge University Press.
Luskin RC, O'Flynn I, Fishkin JS, et al. (2014) Deliberating across Deep Divides. *Political Studies* 62 (2): 116–135.
MacKenzie MK and Warren M (2012) Two Trust-Based Uses of Mini-Publics in Democratic Systems. In: Parkinson J and Mansbridge J (eds) *Deliberative Systems*. Cambridge: Cambridge University Press, pp.52–71.
Macq H and Jacquet V (2023) Institutionalising Participatory and Deliberative Procedures: The Origins of the First Permanent Citizens' Assembly. *European Journal of Political Research* 62: 156–173.
Már K and Gastil J (2020) Tracing the Boundaries of Motivated Reasoning: How Deliberative Minipublics Can Improve Voter Knowledge. *Political Psychology* 41 (1): 107–127.
Martin D and Nai A (2024) Deepening the Rift: Negative Campaigning Rosters Affective Polarization in Multiparty Elections. *Electoral Studies* 87: 102745.
Meléndez C and Kaltwasser CR (2021) Negative Partisanship Towards the Populist Radical Right and Democratic Resilience in Western Europe. *Democratization* 28 (5): 949–969.
Mercier H and Landemore H (2012) Reasoning Is for Arguing: Understanding the Successes and Failures of Deliberation. *Political Psychology* 33 (2): 243–258.
Neblo M, Esterling K and Lazer DMJ (2018) *Politics with the People. Building a Directly Representative Democracy*. Cambridge: Cambridge University Press.
Ober J (1989) *Mass and Elite in Democratic Athens: Rhetoric, Ideology, and the Power of the People*. Princeton, NJ: Princeton University Press.
Orhan YE (2022) The Relationship Between Affective Polarization and Democratic Backsliding: Comparative Evidence. *Democratization* 29 (4): 714–735.
Reiljan A (2020) Fear and Loathing Across Party Lines (also) in Europe: Affective Polarisation in European Party Systems. *European Journal of Political Research* 59 (2): 376–396.
Richey M (2012) Motivated Reasoning in Political Information Processing: The Death Knell of Deliberative Democracy? *Philosophy of the Social Sciences* 42 (4): 511–542.
Rountree J, Anderson C, Reedy J, et al. (2022) The Internal Dynamics of 'Scaling Up' Deliberative Mini-Publics. *Communication and the Public* 7 (3): 146–164.
Rummens S and Geenens R (2023) Lottocracy Versus Democracy. *Res Publica*. Epub ahead of print 22 November. DOI: 10.1007/s11158-023-09648-6.
Schumpeter, J (1962 [1942]) *Capitalism, Socialism and Democracy*. New York: Harper & Row.
Setälä M and Smith G (2018) Deliberative Mini-Publics. In: Bächtiger A, Dryzek J, Mansbridge J, et al. (eds) *The Oxford Handbook of Deliberative Democracy*. Oxford: Oxford University Press, pp.300–314.
Setälä M, Christensen HS, Leino M, et al. (2023) Beyond Polarization and Selective Trust – A Citizens' Jury as a Trusted Source of Information. *Politics* 43 (4): 472–488.
Siu A and Wakao S (2022) Effectiveness of Deliberation Using AI-Assisted Online Platform: Comparative Study of AI-Assisted Online Versus In-Person Deliberative Poll in Japan. In: *Paper presented to the ECPR annual general conference*, *Innsbruck*, 22–26 August.
Slothuus R and de Vreese C (2010) Political Parties, Motivated Reasoning, and Issue Framing Effects. *Journal of Politics* 72 (3): 630–645.
Strickler R (2018) Deliberate with the Enemy? Polarization, Social Identity, and Attitudes Toward Disagreement. *Political Research Quarterly* 71 (1): 3–18.
Taber CS and Lodge M (2006) Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* 50 (3): 755–769.
Wagner M (2024) Affective Polarization in Europe. *European Political Science Review* 16: 378–392.
Warren ME (2020) Participatory Deliberative Democracy in Complex Mass Societies. *Journal of Deliberative Democracy* 16 (2): 81–88.

Warren ME and Gastil J (2015) Can Deliberative Minipublics Address the Cognitive Challenges of Democratic Citizenship. *The Journal of Politics* 77 (2): 582–574.
White A and Neblo N (2021) Capturing the Public: Beyond Technocracy & Populism in the U.S. Administrative State. *Daedalus* 150 (3): 172–187.

Author Biographies

Maija Setälä is a professor in political science at the University of Turku, Finland. In her research, she specializes in democratic theory, democratic innovations and political trust. She has published widely on these topics, including a recent co-edited book *"Democracy and the Future"* (with Michael K. MacKenzie and Simo Kyllönen; EUP 2023).

Ian O'Flynn is Professor of Political Theory in the School of Geography, Politics and Sociology, Newcastle University, UK. His main research interest is in the theory of deliberative democracy, though he also works on topics such as compromise, power-sharing and referendums.