

Opinion

Repairing cognitive distortions in political divides

Luiza A. Santos ^{1,*}, Catherine Garton¹, and Jamil Zaki^{1,*}

Political animosity is on the rise, undermining both well-being and collective responses to global challenges. Interventions to decrease animosity and division often fail to yield lasting, generalized benefits. Drawing on cognitive behavioral therapy (CBT) and new empirical evidence, we argue that entrenched political conflicts are laden with cognitive distortions similar to those observed in clinical contexts. We propose that, by adapting CBT insights to political settings, interventionists can complement and enhance current depolarizing efforts. We provide suggestions for how future efforts should equip individuals to recognize and correct biased thinking outside of the laboratory, offering a pathway toward durably mitigating division and hostility.

Rising threat of political animosity

Political division is straining democracies worldwide. Leaders and legislators have become increasingly divided and prone to incivility, hindering efforts to address global challenges, from climate change to health emergencies [1–4]. These conflicts have seeped into public discourse, shaping how citizens perceive one another. In many instances, people now dislike those who support the opposing side more than they actually disagree with them [4–7]. This rise in hostility is fracturing social ties [8,9], eroding trust [10], intensifying anxiety [11,12], and undermining confidence in democratic governance [13].

In the last quarter century, over 10 000 scientific publications have grappled with political animus and how to mitigate it (for a review, see [14]). While this research has provided valuable insights, it has notable limitations. For instance, interventions to reduce animosity often fail to lessen other corrosive attitudes, such as support for political violence and undemocratic actions [15,16]. The impact of these interventions also tends to be short-lived, with effects fading once individuals return to polarizing environments [14].

Here, we propose that CBT offers a valuable and underutilized framework for mitigating acrimonious division. Extensive work has documented politically biased cognition, showing how partisan identities can skew perception and deepen divides [17,18]. In the first section, we review this work through the lens of CBT. Our aim is to illuminate striking parallels between political biases and clinical cognitive distortions and demonstrate that existing knowledge can be usefully structured within this framework. In the second section, we demonstrate that many promising **depolarization interventions** (see [Glossary](#)) draw, often inadvertently, on CBT principles. We argue that explicitly embracing these principles provides insights about the cyclical mechanisms of intergroup conflict and offers a practical toolkit for improving the efficacy, generalizability, and durability of future efforts.

CBT offers insights for bridging political divides

More than 60 years ago, Aaron Beck proposed that illnesses that appear primarily emotional, such as depression, are often rooted in distortions of thought [19]. According to Beck's cognitive

Highlights

Political animosity has risen, as has interest in mitigating it, resulting in a flood of research and interventions.

The effects of depolarization interventions tend to be circumscribed and short-lived.

We propose that current efforts to reduce ideological division and hostility often fall short because they focus on consequences of cognitive distortions, such as inaccuracy on a particular aspect of polarization, without explicitly addressing the distortions themselves.

Cognitive behavioral therapy (CBT) offers a meaningful framework for understanding divisions and strengthening depolarization efforts.

CBT strategies, such as identifying recurring patterns of distorted thinking, emphasizing skill-building, and fostering sustained contact and support, can enhance the effectiveness of depolarization efforts.

¹Department of Psychology, Stanford University, Stanford, CA, USA

*Correspondence:
lsantos1@stanford.edu (L. A. Santos)
and jzaki@stanford.edu (J. Zaki).



model, exaggerated negative beliefs spawn maladaptive behaviors. For example, if a person believes they are unlovable or inferior, they might withdraw from social activities and opportunities for development. This avoidance, in turn, reinforces the person's sense of inferiority. Instead of focusing on the proximal issue – sadness – Beck and his patients engaged in collaborative empiricism; the practice of jointly identifying and challenging cognitive distortions [20].

We propose that entrenched conflict fosters cognitive distortions which, like those in clinical settings, lead to cycles of counterproductive behavior that feed back into distorted thinking. While many depolarization interventions have successfully challenged specific misbeliefs, such as overestimating how much the outgroup dislikes one's own party, they often fail to address the broader cognitive patterns that produce these errors. For example, learning that political opponents dislike one's own party less than expected may momentarily reduce animosity, but if underlying distortions remain unaddressed, individuals may continue to interpret other outgroup actions through a hostile lens. A CBT framework helps diagnose these underlying patterns and clarifies how existing interventions might be adapted to promote more lasting change.

Parallels between clinical and political cognitive distortions

Beck and others pinpointed common cognitive distortions. For instance, depressed patients think in black-and-white terms, overgeneralize based on limited evidence, magnify the negative, and catastrophize – drawing drastic interpretations of ambiguous events. A wellspring of new evidence suggests that similar distortions characterize cognition in polarized settings (Figure 1 and Box 1).

Black-and-white thinking

Black-and-white thinking involves perceiving situations in extreme and mutually exclusive terms. For example, someone struggling with generalized anxiety disorder may view their performance in a task as either flawless or a complete failure, with no in between. In political contexts, black-and-white thinking manifests as dichotomous views of complex political phenomena. People on both sides of the political spectrum often engage in zero-sum thinking, ascribing any positive outcome to the other side as a loss for their own [21].

Relative to moderates, people who are more politically extreme view politics in simplistic categories [22]. For example, in the context of the EU refugee crisis, people on the extremes of the ideological spectrum were more likely than moderates to cling to simple answers (e.g., letting all refugees in or keeping them all out), and were more certain of their position, irrespective of their subject knowledge [23].

People in polarized environments also tend to moralize political stances, portraying their own side's stereotypical positions as entirely correct and the opposing side's views as wholly immoral, even on issues where there is diverse opinion within each political group [24,25]. Partisans even mistakenly believe those on the other side support blatant wrongs such as child pornography and animal abuse [26]. These beliefs fuel the dehumanization of political outgroups and shape reluctance to interact across political disagreements [26].

Black-and-white thinking can also impede compromise [27]. When people perceive fundamental moral differences between their ingroup and the outgroup, they become more supportive of authoritarian leaders [28] and engage in and justify violent political behavior [29].

Overgeneralization

Overgeneralization is the inclination to draw sweeping conclusions based on limited evidence. For instance, a person with depression might view a setback, such as a friend canceling a date, as an

Glossary

Bridging organizations: nonprofit groups dedicated to overcoming political division.

Collective action: the coordinated effort of a group working toward a shared goal to improve their circumstances.

Depolarizing interventions: structured efforts designed to improve cross-ideological interactions. These interventions can focus on reducing animosity, disagreement, and/or inaccuracies between political groups.

Outpartisan: people who support the opposing political party – used most commonly in two-party systems.

Political elites: a small, influential group of individuals who wield significant power in a society, often shaping policy, governance, and public discourse. This group typically includes high-ranking government officials, legislators, judicial leaders, party strategists, and other key political figures who influence decision-making at national or regional levels.

Socratic questioning: an approach to dialogue that employs open-ended questions to foster critical thinking and challenge assumptions.

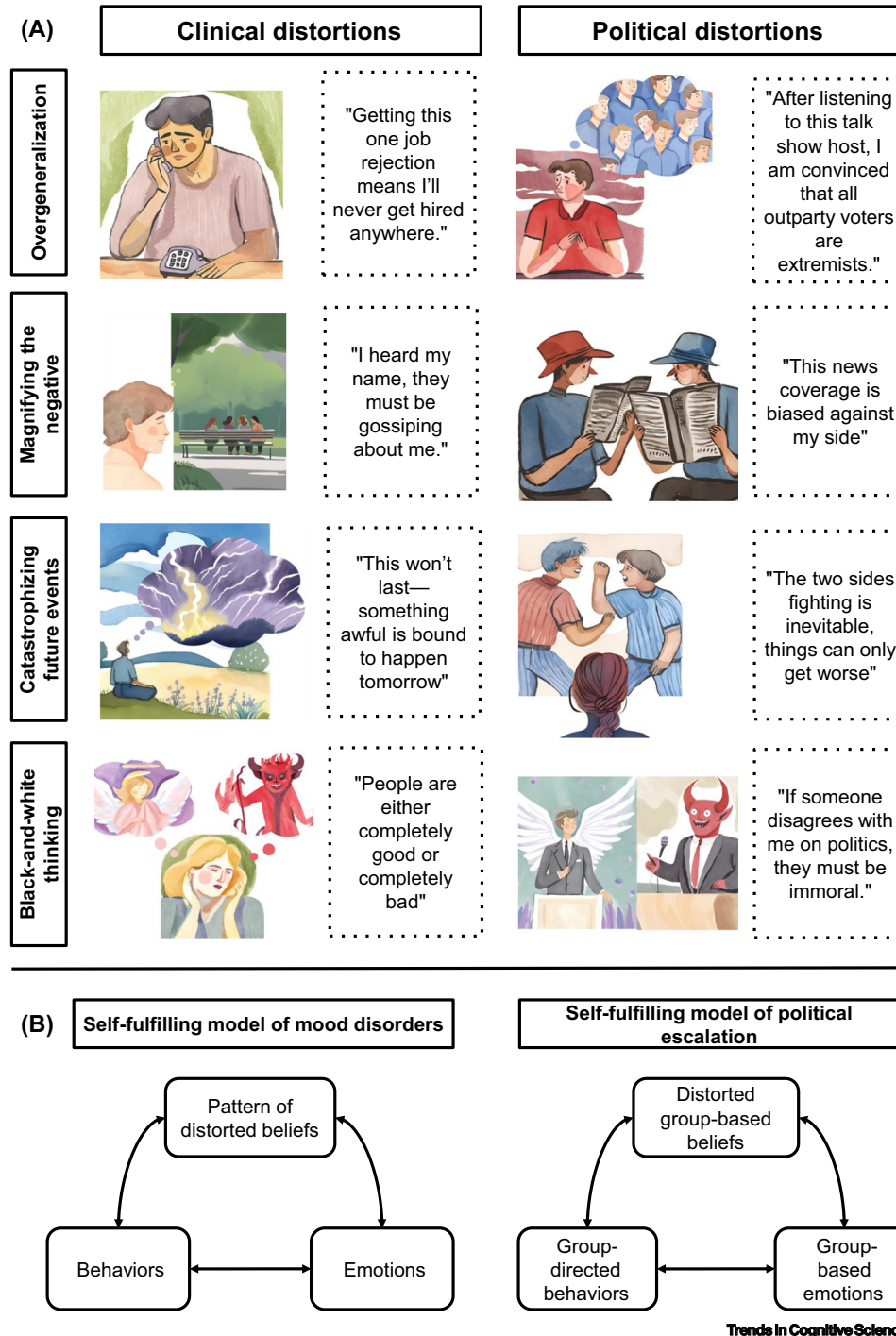


Figure 1. Exemplifying cognitive distortions. (A) Nonexhaustive list of cognitive distortions observed in both clinical and political contexts, with dashed boxes illustrating common beliefs associated with each distortion. (B) Schematic model of self-escalating distortions in both clinical (right) and political (left) settings: erroneous beliefs drive intense emotions and maladaptive behaviors, which, in turn, confirm the original distorted beliefs. All images were created using Canva, a graphic design software.

Box 1. Distinguishing between political cognitive distortions and mental illness

Although we draw parallels between cognitive distortions in political and clinical contexts, we do not equate partisan animosity with mental illness. Mental illnesses are clinically significant disturbances in cognition, emotion regulation, or behavior that cause substantial distress or dysfunction, as defined by diagnostic manuals like the DSM-5 or ICD-11. By contrast, political biases – while they may lead to personal or societal harm – are widespread, often normalized, and not primarily understood as a pathology. There is no public or professional consensus that political animosity is dysfunctional. Our goal is to illustrate how insights from CBT, which targets distorted thought patterns, can inform strategies to reduce partisan division and hostility, without implying that political disagreements or animosity should be categorized as a mental disorder.

indication that they should never rely on others. People also consistently overgeneralize in polarized contexts, making negative assumptions about entire groups.

Across 26 countries, people exaggerate the extremism of political outgroups [30]. For example, in the USA, Democrats and Republicans perceive **outpartisans** as twice as likely to support undemocratic practices and at least three times as likely to support political violence than they actually are [31,32]. Partisans also overestimate the uniformity of outgroup opinions, extending stereotypical positions to all outgroup members [24].

CBT models hold that overgeneralization leads to self-fulfilling prophecies: a depressed patient who believes no one wants to spend time with them might avoid social interactions and lose friendships. Similarly, political misperceptions can lead people to act in ways that confirm their initial fears [33]. For instance, believing that the majority of political rivals are undemocratic or violent leads people to justify and support their own party's undemocratic and violent behavior [31,32].

Catastrophizing future events

Catastrophizing future events is the tendency to anticipate dire outcomes from ambiguous situations. After a bad date, someone suffering from generalized anxiety disorder may conclude that they will never find love. In a polarized environment, individuals may outweigh the likelihood of disastrous political outcomes. For instance, 40% of Americans believe the USA is on the brink of a second civil war¹, even though civil wars – defined as internal conflicts resulting in >1000 battle-related deaths within a year [34] – are exceedingly rare in western democracies. This shows how cognitive distortions can reinforce one another: overgeneralizing the outgroup's propensity for violence can heighten catastrophized fears about the future.

Another type of political catastrophizing is the belief that conflicts are irreparable. In long-term conflicts, those who lack hope for intergroup reconciliation are also the least supportive of peace-promoting policies [35]. Over time, this sense of collective hopelessness becomes ingrained into the national narrative and becomes a core psychological mechanism sustaining division [36]. In other words, when people predict dire and inevitable political conflict, their actions can inadvertently sustain it.

Magnifying the negative

Magnifying the negative involves viewing ambiguous situations through a negative lens. For example, a depressed person might receive neutral feedback at work and interpret it as subtle criticism. In polarized environments, individuals selectively interpret political events in ways that confirm polarization. For instance, people on both sides of the political divide find the same media content unfairly biased against the position they support – a phenomenon known as hostile media bias [37].

Even perceptions of objective economic conditions can be skewed through a partisan lens [38]. In the USA, <30 days after Donald Trump's 2025 inauguration, Democrats' ratings on economic conditions declined by 30 points, while Republicans' assessments improved by the same marginⁱⁱ.

Political identification even shapes selective or false recall of memories. In one recent study, Democrats were more likely to recall fictitious events that seemed damaging to Republicans, such as President Trump allowing companies to dump waste in the Great Lakes. Republicans were more likely to remember false events that would confirm their negative beliefs about Democrats, such as President Obama colluding with liberal media for positive coverage [39].

In clinical settings, this distortion reinforces negative feelings, furthering symptomatology. In political contexts, it makes people believe the worst about their opponents and fortifies polarized identities, even when evidence contradicts their assumptions.

Self-reinforcing cycles in political conflict

CBT posits that distorted beliefs generate maladaptive emotions and behaviors, which in turn reinforce the original beliefs in a self-perpetuating loop. This dynamic extends to political contexts as well (Figure 1B). For example, believing that the opposing party is inherently hostile can evoke anger and prompt aggressive engagement with outgroup members online, interactions that often serve to confirm the initial belief. Over time, these cycles can unfold collectively, intensifying group-based animosity. Although prior research has highlighted self-fulfilling prophecies as barriers to conflict resolution [40], a CBT framework emphasizes the role of cognitive distortions in perpetuating hostility. By applying this perspective, we aim to bridge intergroup conflict research with clinical science and underscore the importance of cultivating robust psychological skills that can disrupt these reinforcing loops.

Factors that fuel political cognitive distortions

Cognitive distortions in polarized settings can originate from various sources, including the proliferation of inflammatory political news, the rapid rise of social media, the growth of polarizing political leadership, a decline in civil discourse, and the increased segregation of the electorate. Once established, these distortions spread rapidly and become self-reinforcing – prompting actions that worsen the very biases from which they stem.

In some countries, the televised media ecosystem is becoming fragmented along ideological lines, leading citizens to inhabit increasingly siloed informational worlds. In the United States, viewers increasingly gravitate toward televised news that aligns with their political views [41] and deeply distrust the outlets relied on by the other sideⁱⁱⁱ.

Likewise, on social media platforms such as X, a small group of ideologically extreme users generates the bulk of political content [42]. This skew in content is compounded by imbalanced sharing patterns: outrage-inducing and moralized posts spread far more widely than balanced political discussions [43–45]. As a result, this online ecosystem paints a misleading picture, making extreme opinions appear far more widespread and representative than they truly are [46,47].

Political elites can also be a source of distorted thinking. Politicians' polarization has accelerated faster than that of the general public [3,48]. However, people often perceive the public as nearly as polarized as elites, suggesting they may generalize elite behavior to society at large [49]. Political elites have also grown more uncivil over time. An analysis of 1.3 million Twitter posts by US Members of Congress (2009–2019) found a 23% rise in incivility over that decade [2]. A study of the 2018 Italian general election also noted a pervasive use of uncivil attacks by candidates online [50]. Political leaders shape norms around acceptable conduct in political disagreements [51]. Across a range of issues and countries, merely invoking messages from political elites can polarize the public's attitudes and increase their animosity [52–54].

Populist rhetoric, which has resurfaced prominently in the 21st century, also amplifies cognitive distortions. Populist leaders frame political debates in terms of a struggle between 'the people' and 'corrupt others', advocate for challenging the establishment, and express antagonism towards ostensible outsiders [55]. Populism normalizes the vilification of opponents and uses crisis-laden narratives that foment black-and-white thinking and catastrophizing [56].

Finally, people are increasingly sorted by political differences. In the USA, people tend to live in politically homogeneous neighborhoods [57] and express a desire to move away from communities that do not share their political views [58]. Even when opportunities to engage across party lines arise, many choose to avoid them. For example, Thanksgiving dinners are shortened, by as much as 50 min, when family members come from opposing voting precincts [8]. As direct contact diminishes, an information vacuum emerges, allowing cognitive distortions to grow.

CBT-inspired treatments for political division

In Beck's hands, CBT was not merely a theory of disorder but also a recipe for new treatments. As such, CBT offers insight into the types of interventions that may be successful in reducing division and hostility. We organize several novel interventions into interrelated categories of CBT techniques: reality testing, cognitive reframing, and behavioral experiments (Table 1).

Table 1. Categories of CBT-relevant treatments and their characteristics

Category of CBT-relevant treatments	Example treatment characteristics	Definition of characteristics	Empirical examples
Reality testing	Misperception correction	The extent to which the treatment corrects misperceptions of outpartisans	[15,26,30–32,82,85–87]
	Heterogeneity of views	The extent to which the treatment displays the heterogeneity of outgroup attitudes or individualizes outgroup members	[24,88]
	Countervailing narratives	The extent to which the treatment challenges traditional depictions of conflict and highlights unity	[15,16,59–61,89]
Cognitive reframing	Emotional reappraisal	The extent to which the treatment helps partisans reframe their emotional experiences	[62,77,83]
	Emotional beliefs	The extent to which the treatment changes people's beliefs about conflict-relevant emotions	[32,63,64,90]
Behavioral experiments	Intergroup contact	The extent to which the treatment fosters intergroup contact	[32,63,65–67,91]
	Dialog training	The extent to which the treatment teaches people better ways to interact across differences	[72,92,93]

Reality testing

Patients in CBT are encouraged to challenge their distorted thoughts through evidence. For instance, if a patient thinks, 'If I make a mistake at work, I'll never succeed again', the therapist might help the patient critically evaluate the validity of their belief by searching for counterevidence.

A growing body of literature on misperception corrections suggests that reality testing can also be beneficial in polarized contexts (for a review, see [33]). In these studies, participants first estimate the attributes of the outgroup – for example, the extent to which the average member of a political outgroup dislikes their side. Then, participants are randomly assigned to receive either accurate information about the outgroup's actual views (known as a misperception correction) or no additional information. Results from these studies demonstrate that misperception corrections can interrupt cycles of escalation. For instance, when people learn that outpartisans are opposed to political violence and undemocratic practices, they become less supportive of these actions themselves [31,32]. Likewise, when participants learn that outgroup members hold more diverse opinions than they initially believed, they perceive the outgroup as less threatening and experience less animosity [24].

Reality testing can also take the form of countervailing narratives – qualitative information that challenges traditional depictions of conflict and highlights unity. For example, witnessing two people with deeply opposing views engage in amicable conversations about their differences reduces partisan animosity among viewers [15]. Similarly, learning about high-profile cross-party friendships (e.g., Joe Biden and John McCain) leads people to feel less negatively toward the opposing side [16]. Countervailing narratives can also help reduce opinion polarization; for instance, Republican voters become more trusting of the electoral process and more supportive of COVID-19 vaccines when they see Republican leaders endorsing these efforts [59,60]. A real-world example of this approach is the Disagree Better Campaign, which ran advertisements depicting US governors from different parties working together respectfully.

Finally, reality testing can challenge catastrophic thinking about the irreparability of conflict. In polarized settings, people can become convinced that groups possess fixed traits, such as believing that outgroups are evil and will always remain that way. This belief precludes the possibility of reconciliation. Israelis who participate in a training program emphasizing that groups are capable of change become more willing to make concessions for peace [61].

Cognitive reframing

The second type of CBT-style treatment pertains to cognitive reframing. This technique helps patients recognize that their initial interpretation of a situation may be overly negative and prompts them to explore different viewpoints. For example, after receiving a bad grade on a test, a patient might believe they are not smart enough to succeed. The therapist might guide the patient to consider how they would respond if a friend had the same reaction to one disappointing exam.

In political contexts, cognitive reframing can reduce negative emotions and increase support for reconciliation. In one experiment, people who were taught to respond to conflict-relevant information like scientists experienced less anger and were less supportive of aggressive political actions [62]. **Bridging organizations**, like Braver Angels and the Constructive Dialogue Institute, use cognitive reframing techniques to help participants adopt an explorer mindset and reappraise stress during disagreements.

Cognitive reframing techniques can also reshape people's beliefs about emotions in political contexts. For example, people often underestimate the practical benefits of empathizing with political

rivals. However, learning about its instrumental value encourages greater empathic engagement across differences, reducing animosity and fostering common ground [63]. Additionally, interventions that frame empathy as an abundant rather than a scarce resource enhance prosocial behavior toward outgroup members [64].

Behavioral experiments

A third category of CBT-relevant treatments encourages people to gather new data. The goal is to help the patient see for themselves that their negative beliefs are often unsupported by actual experience, thereby promoting more adaptive thinking. A patient who fears public speaking might be encouraged to give a short talk in front of friends to see if their prediction of rejection comes true.

In the political realm, conversations across ideological divides can function as powerful behavioral experiments. Rather than relying on secondhand information, direct engagement provides an opportunity to update beliefs about others and their social groups. A person expecting hostility from a political opponent may instead find openness or nuance. For instance, one study found that participants expected conversations with political opponents to be far less positive than discussions with like-minded individuals. However, participants found conversations across political divides to be as informative and enjoyable as those with people who shared their views [65].

In addition to being more pleasant than expected, these conversations can also improve common ground. Across several experiments, conversations across political differences can lead people to durably revise their negative stereotypes about outgroups [66]. Many bridging organizations, such as Living Room Conversations, facilitate workshops that allow constituents to connect face-to-face to explore policy disagreements and identify shared values. Even less personal settings, such as texting anonymously with an outpartisan, can increase common ground on contentious issues, like immigration and gun control [67].

Not all intergroup contact yields positive outcomes [68]. Just as in clinical settings, the effectiveness of behavioral experiments in the political realm relies on creating a safe environment for individuals to test their assumptions. These cross-party contact interventions are more likely to yield positive results when interactions are civil [67], empathic [63], and encourage people to learn more about the perspectives of outgroup members [69].

Improving depolarizing interventions with CBT principles

As we have reviewed, existing depolarizing interventions often connect with principles from CBT. That said, those connections tend to be incidental. We argue that applying CBT more deliberately can strengthen interventions.

First, whereas CBT techniques explicitly help people recognize and correct broader cognitive distortions, depolarizing interventions focus on correcting specific incorrect thoughts one at a time (Figure 2A). This circumscribed approach limits their effectiveness, as correcting one type of inaccuracy does not necessarily lead to improvements in others. For instance, when participants learn that outpartisans dislike them less than they initially believed, they do not become more realistic about outpartisans' support for democracy or peace [15,70]. These interventions also are not robust to the presence of conflicting information, undermining their applicability in real-world scenarios [71]. By applying CBT principles, interventions could teach individuals how these isolated inaccuracies are part of a larger cognitive distortion (Figure 2B). For example, researchers might design a repeated exposure study that teaches participants about common distortions and helps them reflect on whether their exaggerated perceptions of outpartisans' hostility, propensity

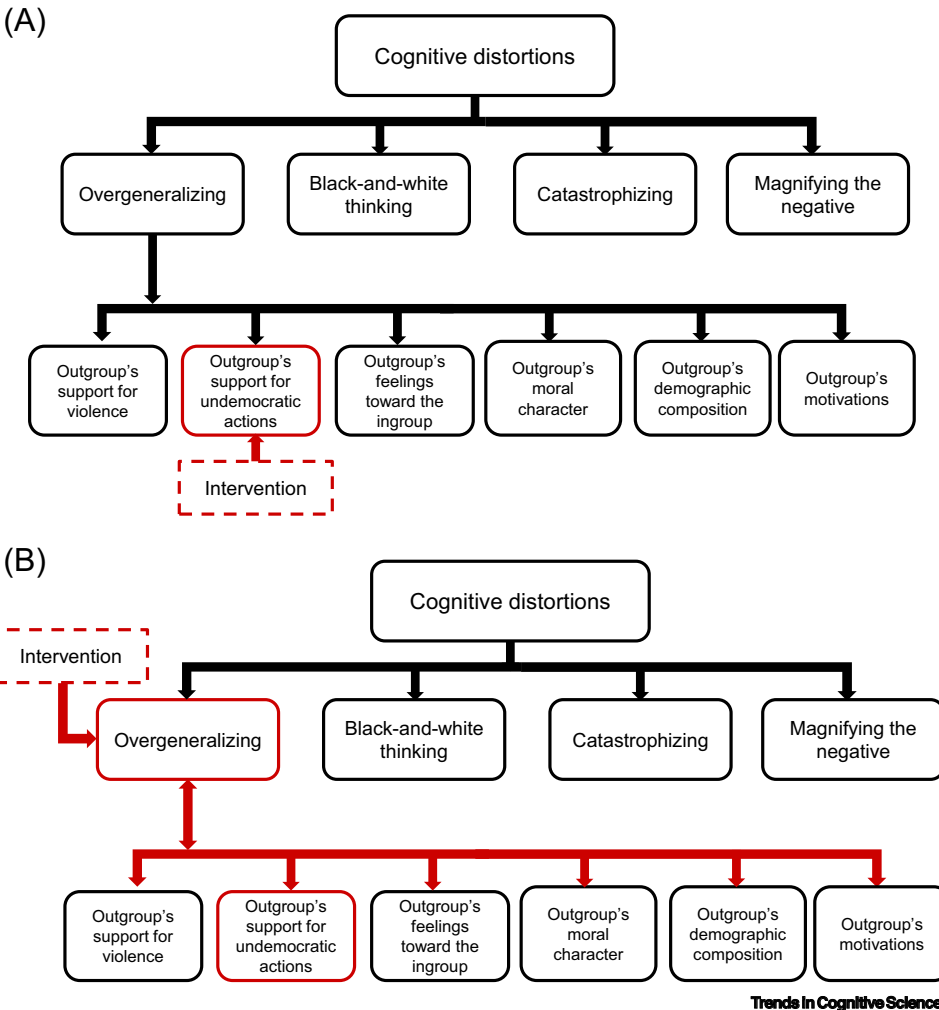


Figure 2. Comparing approaches to addressing political divisions. (A) Example of how existing interventions treat political divisions by correcting one downstream political misperception at a time. (B) Schematic depiction of how treating divisions by helping people identify and challenge upstream cognitive distortions can cast a wider net – potentially triggering cascading corrections across multiple misperceptions. Cognitive behavioral therapy aims for patients to become their own therapists, identifying and challenging distortions on their own. This figure illustrates how depolarization interventions can pursue a similar goal in political contexts.

for violence, and undemocratic support reveal a pattern of overgeneralization. This active learning approach could bolster the generalizability and robustness of the intervention effects.

Second, CBT seeks to enhance patients' self-efficacy by equipping them with concrete skills to identify and correct cognitive distortions in everyday life. For example, patients are taught to write diaries to monitor their beliefs and emotions. They also engage in **Socratic questioning**; a structured method that prompts critical thinking through questions like 'What is the evidence for this thought?' and 'Is there an alternative perspective?'. In contrast, many depolarizing interventions treat participants as passive recipients of information, offering statistics or curated stories without guiding people in questioning their own biases. This approach fails to equip individuals with transferable skills they can apply once they finish a study and re-enter divisive information environments. CBT-inspired interventions should emphasize skill-building, such as

seeking counterexamples to negative assumptions or practicing conversational strategies that improve political discourse [72,73]. For example, researchers could design a daily diary study in which participants reflect on recent political thoughts or conversations and are prompted to identify any cognitive distortions they may have exhibited. Participants could then be guided to generate alternative interpretations and track how these reappraisals influence their emotions and subsequent behavior. By equipping individuals with these tools, such interventions can allow them to self-correct outside of the lab, potentially offering resilience against the pressures of an increasingly polarized society.

Third, we hope that a CBT framework invites scholars to engage more deeply with adaptive thinking strategies developed in clinical psychology. For example, dialectical behavior therapy cultivates dialectical thinking – the ability to hold and integrate seemingly opposing viewpoints – and distress tolerance – the capacity to manage intense emotions without resorting to harmful behaviors. Acceptance and commitment therapy promotes values-driven action over emotional reactivity. These clinically developed techniques are complementary to the aims of depolarization interventions and offer an underexplored toolkit for improving them.

Fourth, CBT illuminates the degree of sustained effort that is often required to achieve long-term change. CBT treatments require an average commitment of 12–20 weeks and rely on a collaborative relationship between patient and therapist [74]. In contrast, depolarizing interventions often rely on quick, light-touch, one-size-fits-all treatments that are increasingly deployed anonymously online. We often expect durable cognitive change from these brief, superficial engagements, overlooking the sustained effort necessary to create lasting shifts in beliefs and behaviors [75].

Caveats to the model

We argue that CBT provides a useful framework for understanding and mitigating political conflict. However, there are some important caveats to a CBT model of division.

First, while CBT is effective in improving affect, negative emotions are not always maladaptive in political contexts. Emotions such as outrage can play a crucial role in motivating **collective action** [11,76]. As such, interventions that reduce animosity might diminish individuals' willingness to engage in political behaviors, such as volunteering or protesting [77]. However, high levels of chronic negative emotions can also decrease collective action by increasing cynicism and hopelessness [78]. We propose that animosity becomes detrimental when it is based on cognitive distortions. When individuals hate political opponents for views they do not actually hold, they risk engaging in behaviors that needlessly escalate conflict. Future work should investigate how to mitigate cognitive distortions while preserving people's motivation for constructive political participation (see [Outstanding questions](#)).

Second, while cognitive distortions may be common across affectively polarized societies, CBT-informed interventions should be culturally adapted. For instance, behavioral experiments involving direct conversations across political divides may be less appropriate in cultures where open disagreement is discouraged. In such contexts, interventions might rely on less confrontational formats, such as structured storytelling exercises. A recent narrative intervention in West Africa leveraged popular radio dramas to reduce support for violence [79], demonstrating the power of culturally resonant formats to promote attitudinal change.

Third, unlike clinical patients seeking relief, politically polarized individuals may not want to change [80], posing an important challenge to the applicability of CBT-informed interventions. Yet, people often experience physical and emotional costs from exposure to political divisiveness [12],

creating an opportunity to design interventions that align with motivations such as self-improvement and enhanced well-being. CBT offers several strategies for engaging ambivalent patients. Motivational interviewing, for example, helps individuals clarify personal goals – such as reconnecting with others or reducing conflict – and examine how their current beliefs might undermine those aspirations. Researchers can also draw on entertainment-based strategies: games are widely used in CBT to promote cognitive restructuring, and depolarizing games have shown promise in reducing animosity, with some effects lasting up to 4 months [81]. These strategies can support both recruitment and retention, helping to scale depolarization efforts.

Fourth, whereas our approach focuses on individual-level strategies to mitigate cognitive distortions, these distortions often stem from complex systemic structures. Although CBT is typically applied at the individual level, it lends theoretical support for system-level interventions. Examples include social media ranking algorithms that incentivize constructive engagement, interventions targeted at political leaders [82], and integrating CBT skill-building into educational curricula and public campaigns. We believe that public awareness itself can be a mechanism of change. In the clinical domain, recognition of cognitive distortions in mood disorders has contributed to increased openness to therapy. In the political context, knowledge about common distortions could normalize the idea that these beliefs should be challenged. CBT-inspired interventions can also address group-level conflict by harnessing social network dynamics, where a single individual's change can ripple through their community. For example, in the Israeli–Palestinian context, an emotion regulation intervention with a subset of group members benefited the rest of the group [83], and teaching one person empathy strategies reduced partisan animosity among their outgroup contacts [63]. These effects could be amplified by targeting highly connected individuals, such as political elites, media figures, or community leaders, whose influence can shift broader group norms [84].

Concluding remarks

The latest World Economic Forum report listed political polarization as one of the top global risks facing the world today^{iv}. There are real shared threats in our political worlds, and effectively addressing these threats requires clarity. Cognitive distortions obscure this clarity and divert energy toward battling caricatured versions of political opponents. By drawing from CBT, we offer a framework for understanding how these distortions escalate conflict and highlight how current interventions can be enhanced to achieve a broader, more lasting impact.

Declaration of interests

No interests are declared.

Resources

ⁱ<https://today.yougov.com/politics/articles/43553-two-in-five-americans-civil-war-somewhat-likely>

ⁱⁱ<https://news.gallup.com/poll/655859/economic-confidence-ticks-down-partisans-views-shift.aspx>

ⁱⁱⁱwww.pewresearch.org/journalism/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided/

^{iv}www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf

References

1. Druckman, J.N. et al. (2021) How affective polarization shapes Americans' political beliefs: a study of response to the COVID-19 pandemic. *J. Exp. Polit. Sci.* 8, 223–234
2. Frimer, J.A. et al. (2023) Incivility is rising among American politicians on Twitter. *Soc. Psychol. Personal. Sci.* 14, 259–269
3. Vegetti, F. (2019) The political nature of ideological polarization: the case of Hungary. *Ann. Am. Acad. Polit. Soc. Sci.* 681, 78–96
4. Bolsen, T. and Druckman, J.N. (2018) Do partisanship and politicization undermine the impact of a scientific consensus message about climate change? *Group Process. Intergroup Relat.* 21, 389–402
5. Carothers, T. and O'Donohue, A. (2019) *Democracies Divided: The Global Challenge of Political Polarization*, Brookings Institution Press
6. Mason, L. (2018) *Uncivil Agreement: How Politics Became Our Identity*, University of Chicago Press
7. Reiljan, A. (2020) 'Fear and loathing across party lines' (also) in Europe: affective polarisation in European party systems. *Eur. J. Polit. Res.* 59, 376–396

Outstanding questions

How can we effectively balance the need to mitigate cognitive distortions in political discourse while preserving the motivation for meaningful political action?

How can the principles of CBT be scaled to address the roots of cognitive distortions at a societal level?

What mechanisms underlie the way individual cognitive distortions (such as overgeneralization and catastrophizing) reinforce one another, and what strategies can disrupt these feedback loops to mitigate escalation?

Beyond those discussed, how do other distortions, such as emotional reasoning (i.e., assuming emotions reflect objective truth) and the minimization of positive information, contribute to political polarization, and how can they be effectively targeted by interventions?

What cognitive skills are most effective at helping individuals resist future polarization triggers, such as exposure to partisan media or inflammatory rhetoric, and how can these be systematically taught?

How can depolarization interventions ensure that individuals apply cognitive techniques beyond controlled settings, integrating these skills into everyday political discussions and decision-making?

8. Chen, M.K. and Rohla, R. (2018) The effect of partisanship and political advertising on close family ties. *Science* 360, 1020–1024
9. Facciani, M. and McKay, T. (2022) Network loss following the 2016 Presidential Election among LGBTQ+ adults. *Appl. Netw. Sci.* 7, 41
10. Lee, A.H-Y. (2021) How the politicization of everyday activities affects the public sphere: the effects of partisan stereotypes on cross-cutting interactions. *Polit. Commun.* 38, 499–518
11. Ford, B. et al. (2023) The political is personal: the costs of daily politics. *J. Pers. Soc. Psychol.* 112, 357–360
12. Fraser, T. et al. (2022) The harmful effects of partisan polarization on health. *PNAS Nexus* 1, pgac011
13. Kingzette, J. et al. (2021) How affective polarization undermines support for democratic norms. *Public Opin. Q.* 85, 663–677
14. Hartman, R. et al. (2022) Interventions to reduce partisan animosity. *Nat. Hum. Behav.* 6, 1194–1205
15. Voelkel, J.G. et al. (2024) Megastudy testing 25 treatments to reduce antidemocratic attitudes and partisan animosity. *Science* 386, eadh4764
16. Voelkel, J.G. et al. (2022) Interventions reducing affective polarization do not necessarily improve anti-democratic attitudes. *Nat. Hum. Behav.* 7, 55–64
17. Van Bavel, J.J. and Pereira, A. (2018) The partisan brain: an identity-based model of political belief. *Trends Cogn. Sci.* 22, 213–224
18. Robinson, R.J. et al. (1995) Actual versus assumed differences in construal: “naive realism” in intergroup perception and conflict. *J. Pers. Soc. Psychol.* 68, 404–417
19. Beck, A.T. (1963) Idiosyncratic content and cognitive distortions. *Arch. Gen. Psychiatry* 9, 324–333
20. Dattilo, F.M. and Hanna, M.A. (2012) Collaboration in cognitive-behavioral therapy. *J. Clin. Psychol.* 68, 146–158
21. Roberts, R. and Davidai, S. (2022) The psychology of asymmetric zero-sum beliefs. *J. Pers. Soc. Psychol.* 123, 559–575
22. Lammers, J. et al. (2017) The political domain appears simpler to the politically extreme than to political moderates. *Soc. Psychol. Personal. Sci.* 8, 612–622
23. Van Prooijen, J-W. et al. (2018) Ideological responses to the EU refugee crisis: the left, the right, and the extremes. *Soc. Psychol. Personal. Sci.* 9, 143–150
24. Dias, N.C. et al. (2024) American partisans vastly under-estimate the diversity of other partisans’ policy attitudes. *PSRM* 1–11
25. Finkel, E.J. et al. (2020) Political sectarianism in America. *Science* 370, 533–536
26. Puryear, C. et al. (2024) People believe political opponents accept blatant moral wrongs, fueling partisan divides. *PNAS Nexus* 3, 244
27. Kodapanakkal, R. et al. (2022) Moral frames are persuasive and moralize attitudes; nonmoral frames are persuasive and demoralize attitudes. *Psych. Sci.* 33, 433–449
28. Crimston, C.R. et al. (2022) Moral polarization predicts support for authoritarian and progressive strong leaders via the perceived breakdown of society. *Polit. Psychol.* 43, 671–691
29. Mooijman, M. et al. (2018) Moralization in social networks and the emergence of violence during protests. *Nat. Hum. Behav.* 2, 389–396
30. Ruggeri, K. et al. (2021) The general fault in our fault lines. *Nat. Hum. Behav.* 5, 1369–1380
31. Mernyk, J.S. et al. (2022) Correcting inaccurate metaperceptions reduces Americans’ support for partisan violence. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2116851119
32. Santos, L.A. et al. (2024) Positive beliefs about cross-partisan empathy can strengthen Americans’ support for democracy. *PNAS Nexus* 3, pgae394
33. Moore-Berg, S.L. and Hameiri, B. (2024) Improving intergroup relations with meta-perception correction interventions. *Trends Cogn. Sci.* 28, 190–192
34. Blattman, C. and Miguel, E. (2010) Civil war. *J. Econ. Lit.* 48, 3–57
35. Hasler, B.S. et al. (2023) Young generations’ hopelessness perpetuates long-term conflicts. *Sci. Rep.* 13, 4926
36. Bar-Tal, D. (2007) Sociopsychological foundations of intractable conflicts. *Am. Behav. Sci.* 50, 1430–1453
37. Vallone, R.P. et al. (1985) The hostile media phenomenon: biased perception and perceptions of media bias in coverage of the Beirut massacre. *J. Pers. Soc. Psychol.* 49, 577–585
38. Enns, P.K. and McAvoy, G.E. (2012) The role of partisanship in aggregate opinion. *Polit. Behav.* 34, 627–651
39. Armaly, M.T. and Enders, A.M. (2023) Filling in the gaps: false memories and partisan bias. *Polit. Psychol.* 44, 281–299
40. Friend, W. and Malhotra, D. (2019) Psychological barriers to resolving intergroup conflict: an extensive review and consolidation of the literature. *Negot. J.* 35, 407–442
41. Muise, D. et al. (2022) Quantifying partisan news diets in Web and TV audiences. *Sci. Adv.* 8, eabn0083
42. Bor, A. and Petersen, M.B. (2022) The psychology of online political hostility: a comprehensive, cross-national test of the mismatch hypothesis. *Am. Polit. Sci. Rev.* 116, 1–18
43. Brady, W.J. et al. (2021) How social learning amplifies moral outrage expression in online social networks. *Sci. Adv.* 7, eabe5641
44. Brady, W.J. et al. (2017) Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl. Acad. Sci. U. S. A.* 114, 7313–7318
45. Rathje, S. et al. (2021) Out-group animosity drives engagement on social media. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2024292118
46. Lee, A.Y. et al. (2024) People overestimate how many social media users post harmful content. *PsyArchXiv*, Published online October 17, 2024. <https://doi.org/10.31234/osf.io/f8y62>
47. Robertson, C.E. et al. (2024) Inside the funhouse mirror factory: how social media distorts perceptions of norms. *Curr. Opin. Psychol.* 60, 101918
48. Singer, M. (2016) Elite polarization and the electoral impact of left-right placements: evidence from Latin America, 1995–2009. *Lat. Am. Res. Rev.* 51, 174–194
49. Ploger, G. (2024) Polarization all the way down: how coverage of elite and partisan polarization spills over to perceptions of the U.S. mass public. *Polit. Commun.* 41, 393–412
50. Rega, R. and Marchetti, R. (2021) The strategic use of incivility in contemporary politics. The case of the 2018 Italian general election on Facebook. *Commun. Rev.* 24, 107–132
51. Huddy, L. and Yair, O. (2021) Reducing affective polarization: warm group relations or policy compromise? *Polit. Psychol.* 42, 291–309
52. Flores, A. et al. (2022) Politicians polarize and experts depolarize public support for COVID-19 management policies across countries. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2117543119
53. Druckman, J.N. et al. (2013) How elite partisan polarization affects public opinion formation. *Am. Polit. Sci. Rev.* 107, 57–79
54. Bäck, H. et al. (2023) Elite communication and affective polarization among voters. *Elect. Stud.* 84, 102639
55. Mudde, C. and Rovira Kaltwasser, C. (2017) *Populism: A Very Short Introduction*, Oxford University Press
56. Hamelers, M. (2020) They are selling themselves out to the enemy! The content and effects of populist conspiracy theories. *Int. J. Public Opin. Res.* 33,
57. Brown, J.R. and Enos, R.D. (2021) The measurement of partisan sorting for 180 million voters. *Nat. Hum. Behav.* 5, 998–1008
58. Motyl, M. et al. (2014) How ideological migration geographically segregates groups. *J. Exp. Soc. Psychol.* 51, 1–14
59. Clayton, K. and Willer, R. (2023) Endorsements from Republican politicians can increase confidence in U.S. elections. *Res. Polit.* 10, 1–5
60. Pink, S.L. et al. (2021) Elite party cues increase vaccination intentions among Republicans. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2106559118
61. Goldenberg, A. et al. (2018) Testing the impact and durability of a group malleability intervention in the context of the Israeli–Palestinian conflict. *Proc. Natl. Acad. Sci.* 115, 696–701
62. Halperin, E. et al. (2014) Emotion regulation and the cultivation of political tolerance: searching for a new track for intervention. *J. Confl. Resolut.* 58, 1110–1138
63. Santos, L.A. et al. (2022) Belief in the utility of cross-partisan empathy reduces partisan animosity and facilitates political persuasion. *Psychol. Sci.* 33, 1557–1573
64. Hasson, Y. et al. (2022) Using performance art to promote intergroup prosociality by cultivating the belief that empathy is unlim- ited. *Nat. Commun.* 13, 7786
65. Wald, K.A. et al. (2024) Misplaced divides? Discussing political disagreement with strangers can be unexpectedly positive. *Psychol. Sci.* 35, 471–488
66. Kalla, J.L. and Brookman, D.E. (2020) Reducing exclusionary attitudes through interpersonal conversation: evidence from three field experiments. *Am. Polit. Sci. Rev.* 114, 410–425

67. Combs, A. *et al.* (2023) Reducing political polarization in the United States with a mobile chat platform. *Nat. Hum. Behav.* 7, 1454–1461
68. Santoro, E. and Broockman, D.E. (2022) The promise and pitfalls of cross-partisan conversations for reducing affective polarization: evidence from randomized experiments. *Sci. Adv.* 8, eabn5515
69. Kalla, J.L. and Broockman, D.E. (2023) Which narrative strategies durably reduce prejudice? Evidence from field and survey experiments supporting the efficacy of perspective-getting. *Am. J. Polit. Sci.* 67, 185–204
70. Dias, N.C. *et al.* (2024) Correcting misperceptions of partisan opponents is not effective at treating democratic ills. *PNAS Nexus* 3, pgae304
71. Druckman, J.N. (2023) Correcting misperceptions of the other political party does not robustly reduce support for undemocratic practices or partisan violence. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2308938120
72. Itzhakov, G. *et al.* (2024) Listening to understand: the role of high-quality listening on speakers' attitude depolarization during disagreements. *J. Pers. Soc. Psychol.* 126, 213–239
73. Minson, J.A. and Chen, F.S. (2022) Receptiveness to opposing views: conceptualization and integrative review. *Personal. Soc. Psychol. Rev.* 26, 93–111
74. Beck, J.S. and Beck, A.T. (2021) *Cognitive Behavior Therapy: Basics and Beyond*, Guilford Publications
75. Paluck, E.L. *et al.* (2021) Prejudice reduction: progress and challenges. *Annu. Rev. Psychol.* 72, 533–560
76. Spring, V.L. *et al.* (2018) The upside of outrage. *Trends Cogn. Sci.* 22, 1067–1069
77. Ford, B.Q. *et al.* (2019) Using reappraisal to regulate negative emotion after the 2016 U.S. Presidential election: does emotion regulation trump political action? *J. Pers. Soc. Psychol.* 117, 998–1015
78. Chen, C.W. and Gorski, P.C. (2015) Burnout in social justice and human rights activists: symptoms, causes and implications. *J. Hum. Rights Pract.* 7, 366–390
79. Bilali, R. (2022) Fighting violent extremism with narrative intervention: evidence from a field experiment in West Africa. *Psychol. Sci.* 33, 184–195
80. Landry, A. and Halperin, E. (2023) Intergroup psychological interventions: the motivational challenge. *Am. Psychol.* 80, 206–219
81. Woodley, L. *et al.* (2024) Defusing political animosity in the United States with a cooperative online quiz game. *Nat. Hum. Behav.*, Published online June 3, 2025. <https://doi.org/10.31219/osf.io/muyp9>
82. Druckman, J.N. *et al.* (2023) Correcting misperceptions of out-partisans decreases American legislators' support for undemocratic practices. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2301836120
83. Pinus, M. *et al.* (2025) Emotion regulation contagion drives reduction in negative intergroup emotions. *Nat. Commun.* 16, 1387
84. Paluck, E.L. and Shepherd, H. (2012) The salience of social referents: a field experiment on collective norms and harassment behavior in a school social network. *J. Pers. Soc. Psychol.* 103, 899–915
85. Braley, A. *et al.* (2023) Why voters who value democracy participate in democratic backsliding. *Nat. Hum. Behav.* 7, 1282–1293
86. Lees, J. and Cikara, M. (2020) Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts. *Nat. Hum. Behav.* 4, 279–286
87. Landry, A.P. *et al.* (2023) Reducing explicit blatant dehumanization by correcting exaggerated meta-perceptions. *Soc. Psychol. Personal. Sci.* 14, 407–418
88. Koetke, J. *et al.* (2023) Bridging the divide: the effect of individualizing information on attitudes toward political outgroup members. *J. Soc. Polit. Psychol.* 11, 534–554
89. Halperin, E. *et al.* (2011) Promoting the Middle East Peace Process by changing beliefs about group malleability. *Science* 333, 1767–1769
90. Schumann, K. *et al.* (2014) Addressing the empathy deficit: beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *J. Pers. Soc. Psychol.* 107, 475–493
91. Greene, K.F. *et al.* (2025) Interacting as equals reduces partisan polarization in Mexico. *Nat. Hum. Behav.* 9, 147–155
92. Hussein, M.A. and Tormala, Z.L. (2024) You versus we: how pronoun use shapes perceptions of receptiveness. *J. Exp. Soc. Psychol.* 110, 104555
93. Yeomans, M. *et al.* (2020) Conversational receptiveness: improving engagement with opposing views. *Organ. Behav. Hum. Decis. Process.* 160, 131–148