British Journal of Political Science (2025), **55**, e121, 1–22
doi:10.1017/S0007123425100562

British Journal of Political Science

ARTICLE

Deepening, Bridging, and Moving Minds in Stressful Times

Simon Stocker[1], André Bächtiger[1], Bernhard Kittel[2] and Marco Steenbergen[3]

[1]Department of Political Theory and Empirical Democracy Research, University of Stuttgart, Stuttgart, Germany, [2]Department of Economic Sociology, University of Vienna, Vienna, Austria and [3]Department of Political Science, University of Zurich, Zurich, Switzerland
Corresponding author: Simon Stocker; Email: simon.stocker@sowi.uni-stuttgart.de

(Received 3 July 2024; revised 20 May 2025; accepted 20 May 2025)

Abstract
This article focuses on the effects of different *communication modes* – 'contestatory', 'collaborative', and 'open' (and two control groups – 'information-only' and a 'placebo' group) on reasoning and opinion formation in the context of conflictive collective decision problems. Focusing on two population-based survey experiments in Germany and Austria on the prioritization of health or freedom (Germany) and the introduction of mandatory vaccination (Austria) in the COVID-19 crisis, we find an important trade-off: while a contestatory and open mode enhances in-depth reasoning, a collaborative mode promotes constructive thinking. Regarding opinion formation, we find that when societal polarization is not extreme, communication modes do not matter for opinion (de-)polarization; here, the exposure to information is all that is needed to move minds. In highly polarized situations, however, open communication is the only way to communicatively reach out to people. Our results contradict both advocates of a contestatory and collaborative renewal of public discourse.

Keywords: political communication; deliberation; communication modes; complex thinking; opinion change

Introduction

How must we structure our political communications – especially in the increasingly important online sphere – in order to stimulate in-depth as well as constructive thinking about common affairs while at the same time promoting open-mindedness, including opinion change and moderation? This question has gained dramatic importance in current times, where we witness declining levels of argumentative complexity and respectful listening to the 'other side', often in combination with the polarization of opinions (see Dryzek et al. 2019). Several researchers and analysts have argued that we need to 'talk' or 'deliberate' more and think through the bitter disagreements that divide our societies. Indeed, several studies have shown that letting people deliberate together changes minds and depolarizes polarized attitudes substantially (Fishkin et al. 2021; Levendusky and Stecula 2021) as well as enhances the factual knowledge of participants on policy issues (Grönlund et al. 2015). But it is unclear how 'deliberation' – usually implemented as one 'grand design' involving information and deliberation in the form of an open communication mode stressing mutual reason-giving and listening – exactly produces these changes. In this article, we go beyond traditional deliberative exercises and disentangle 'deliberation' into *different communication modes* – 'contestatory', 'collaborative', and 'open'. We study the effects of these

communicative modes on in-depth and constructive thinking as well as changing and (de-) polarizing opinions in the context of two large population-based survey experiments, one with 2,132 respondents in Germany (March 2021) on the prioritization of health or freedom in the COVID-19 crisis and another one on vaccination duty with 2,134 respondents in Austria (November and December 2021). The COVID-19 pandemic is a particularly intriguing topic not only because it affected everyone but also because there were phases with different levels of societal polarization (see, for example, Kerr et al. 2021; Gadarian et al. 2021). The first experiment was conducted when the pandemic entered its second year, where the question of prioritizing health or liberty was contested, but opinions were not fully polarized; the second experiment was conducted about half a year later when societal polarization was particularly high, especially in the case of mandatory vaccination in Austria (see below).

The topic of different communication modes – contestatory v. collaborative – is popular in current debates about democracy's vitality and viability, but it is also theoretically underspecified and empirically understudied. On the one hand, there are several calls to (re-) engage persons with highly divergent political views via proper 'debating' (Manin 2005, 2017; Deitelhoff 2018) and there is also an empirical and formal literature showing that 'adversarial debate' may potentially deepen minds and promote epistemic advancement (Schweiger et al. 1986; Chung and Duggan 2020; Goodin and Spiekermann 2018). On the other hand, there are also serious doubts whether 'adversarial debate' simultaneously facilitates the bridging and integration of conflicting perspectives. A recent public experiment in Taiwan on the strongly disputed regulation of Uber taxis ('vTaiwan') shows that constructive dialogue (where participants focused on potential common ground rather than highlighting differences) conduced to a viable regulatory solution for the involved stakeholders. Prominent deliberative exercises (especially deliberative polling), in turn, involve an open discussion mode (neither debate-style nor explicitly geared toward finding common ground), and this usually conduces to learning and opinion change (see Fishkin 2025). To date, we lack a combined and especially causal test which communication mode – contestatory, collaborative or open – helps to deepen, bridge and move minds (a partial exception is a pioneering study by Schweiger et al. (1986) which, however, focused on epistemic advancement and decision acceptance, but not opinion change).

Our newly developed experimental design first exposed participants to balanced arguments for and against health v. freedom or vaccination duty, and then randomly assigned them to three treatment groups – a 'contestatory', a 'collaborative', and an 'open' communication mode – and to two control groups – 'information-only' and a 'placebo' group (which read a general text on the COVID-19 crisis). We applied minimal communicative interventions: in the 'contestatory' treatment, participants were confronted with a counter-position to their initially stated opinion and were asked to react to this challenge. In the 'collaborative'-treatment, they were asked to combine the two positions (health and freedom or pro- and anti-vaccination duty). In the 'open communication'-treatment, participants were asked to say something about the two positions. By keeping the communicative interventions minimal (with participants only being asked to react to these interventions by writing a short comment), we are in a position to test the communicative modes in a 'pure' way. Yet, one might rightly ask what these minimal interventions tell us about real-world political discourse and its effects on reasoning and opinion change. First, the external validity of experimental interventions is frequently questioned. However, the focus here is on the test of a particular social mechanism which, if its presence is shown in the experiment, can subsequently be explored under more complex conditions. Moreover, since the experiment is embedded in a survey that includes a diverse sample of the target population, the potential objection that the results apply only to an arbitrary group does not apply. Second, and perhaps more importantly, our treatments – focusing on individualist responses to communication modes – cover an important aspect of 'deliberation within' in large mass-scale societies (Goodin 2000) and give novel answers how different modes of political discussion affect our way of thinking and opinion formation (or what Goodin calls the 'internal-reflective' mode). They also align with – and especially expand on – existing interventions in the increasingly important online

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

space, where content moderation – such as empathy-based counter-speech (Hangartner et al. 2021), AI 'conversation-enhancing' interventions (Argyle et al. 2023) or different AI conversational styles (Chen et al. 2024) – have been devised to create more constructive and inclusive online discourses. To our knowledge, our study is the first to experimentally vary such 'prompts' in the context of fundamental communication modes, namely contestatory, collaborative, and open ones, and test relevant effects on in-depth reasoning, constructive thinking and opinion change.

Our results display a complex but intriguing picture. On the one hand, we find stable effects of communication modes on *deepening* and *bridging minds* across the two experiments: both the contestatory and the open mode produced higher levels of in-depth reasoning (measured via justification rationality) while the collaborative mode produced more integration of conflictive perspectives (measured on the basis of constructive proposals). The picture is more complex and fragmented when it comes to *moving* minds. Here, much depends on the level of societal polarization in combination with individual positioning in the societal debate. An intriguing finding is that in the context of a high level of societal polarization, the only mode that produces opinion change (and depolarization) is open communication, where people are just asked to say something about the topic. These results have important ramifications for how we optimally structure our public conversations on matters of common concern in stressful times. Rather than promoting more (adversarial) debate – as Manin (2005) and Deitelhoff (2018) have forcefully argued – or set our hopes into more collaborative discussion modes – as some feminist scholars (Tannen 1998), many discussion and negotiation practitioners (Fisher et al. 1991; Asif and Klein 2009) as well as AI developers (Argyle et al. 2023) have proposed – our results suggest that no communication mode constitutes a 'theory of all seasons' and promotes epistemic advancement, the identification of common ground and moderation *at once*. Instead, we find an important *trade-off* between more in-depth reasoning (produced by the contestatory and open mode) and more constructive and integrative reasoning (produced by the collaborative mode).

The article is structured as follows: First, we formulate expectations about the extent to which different communication modes have effects on factors such as in-depth reasoning, constructive thinking and opinion change (including polarization and depolarization trends). We then introduce the experimental setup, discuss the operationalization of the different variables and present the results of Study 1. Subsequently, we describe Study 2, which was identical in setup and operationalization of variables but focused on the introduction of mandatory vaccination in Austria. Finally, the article concludes with a discussion of the results.

Deepening, Bridging, and Moving Minds in Different Communication Modes

The question of how to deepen, bridge and move minds has been studied from various angles, including cognitive and social psychology, public opinion research and deliberative democracy. In particular, empirical deliberation research has provided rich accounts of how citizens in deliberative events learn, understand the other side, find 'consensus', and both change and depolarize minds (see, for example, Fishkin et al. 2021; Esterling et al. 2018; for an overview, see Bächtiger and Parkinson 2019, Chapter 3). However, one deficiency of existing empirical deliberative research is that it frequently fails to disentangle the exact 'deliberation' effect from other effects (for example, information provision or expert input); moreover, only few studies have begun to analyse which deliberative component – justification rationality, reciprocity or respect – produces these outcomes (an exception is Himmelroos and Christensen 2014). Another important deficiency is the missing focus on different *communication modes* and their effects. Indeed, communication modes may have a considerable independent effect on whether people engage in more in-depth or integrative thinking or change and (de-)polarize minds (see Schweiger et al. 1986). But the topic of different communication modes has barely been picked up by the

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

systematic empirical literature. A partial exception is Ugarriza and Trujillo-Orrego (2020), who examined the effect of different *protocols* – argumentation, free talk and perspective – on attitude changes in small group discussions and showed that only the promotion of perspective taking contributes to reconciliation (an argument also put forward by Muradova and Arceneaux 2022). Another exception is the study by Groenendyk and Krupnikov (2021), who find in an experiment that when communication is 'framed' as debate, people look for arguments that support their previous views; by contrast, when communication is 'framed' as deliberation, people seek consensus and are open to changing their minds. Overall, by disentangling generic notions of 'deliberation' into different communication modes, we also try to re-connect deliberative research with communication studies (Carcasson et al. 2010): while different communication modes are close to the realistic (albeit multifaceted) ways in which political communication proceeds in political practice, they can be linked to deliberative quality standards (such as justification rationality and constructive thinking).

In the following, we focus on three modes of communication: contestatory, collaborative, and open (in connection with control conditions, an information-only and a true control group) and their effects on deepening, bridging and moving minds. The contestatory mode is modeled as 'debate' where participants are confronted with a counter-position and are asked to respond; the collaborative mode is modeled as 'constructive dialogue' where participants engage in 'appreciative inquiry' (Asif and Klein 2009) and are asked to combine conflicting positions; the open mode is modeled as discussion where participants are just asked to say something on the conflicting positions.

Regarding communication modes and *epistemic advancement*, Manin (2017, 2005; Bächtiger 2025) has argued that it is adversarial debate and the 'confrontation of conflicting arguments', rather than constructive dialogue with a focus on common ground and 'interactive' (or open) discussion, that improves the epistemic quality of reasoning (a similar argument has also been made by Goodin and Spiekermann (2018) in the context of the Condorcet Jury Theorem). By the epistemic quality of reasoning, we understand the 'depth of argumentation', focusing on how thoroughly a claim or argument is justified (Bächtiger 2025). The mechanism for the epistemic superiority of contestatory modes is quite straightforward: being confronted with a counterargument or counter-position forces participants to think through their own position, and in so doing, they overcome the usual 'laziness' in reasoning (Mercier 2016) and formulate more in-depth justifications for their positions. As Mercier (2016, 691) puts it, reasoning quality is at its best 'when reasoners devote as much time as possible to evaluating others' arguments, rather than only producing their own'. This also aligns with 'motivated reasoning' approaches (Taber and Lodge 2006) where people confronted with incongruent information 'try to seek out rationales confirming their priors, engaging in effortful cognitive processing' (Erisen et al. 2018, 221).

By contrast, when people are asked to find common ground (as in collaborative modes), this might undermine the production of in-depth justifications. One reason is a simple 'transaction cost' problem: 'finding common ground usually requires that people invest time to identify shared positions and reasons; but this investment may be inversely related to the amount of time that one can invest in finding thorough reasons for various positions' (Bächtiger 2025). Regarding reasoning quality, we expect the open mode to lie between the contestatory and the collaborative mode. Being asked for one's opinion on a topic in an open manner also conduces to justification, although we expect the depth of justification to be lower compared to a contestatory mode where one is directly confronted with counterarguments or counter-positions, which we expect to generate a stronger pressure for producing more in-depth justifications. Chung and Duggan (2020), furthermore, present a formal theory of democratic deliberation, distinguishing among myopic discussion, constructive discussion and debate. They demonstrate formally that debate produces superior outcomes in terms of conclusiveness (that is, outcomes converge on a single position) and non-arbitrariness compared to constructive discussion and especially myopic discussion. While conclusiveness is not the focus of our study, the formal argument by

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

Chung and Duggan (2020) nonetheless suggests that a debate mode might produce a higher quality of reasoning.

There is also some empirical support for these claims. In a pioneering laboratory study, Schweiger et al. (1986) compared the effectiveness of contestatory discussion modes – dialectical inquiry and devil's advocacy – v. consensus approaches in group decision making. Results showed that both dialectical inquiry and devil's advocacy produced a higher level of critical evaluation of assumptions and higher quality recommendations than the consensus treatment. In the context of the Australian Citizen parliament, Curato et al. (2013) find that 'appreciative inquiry' – a collaborative mode with the goal to identify strengths of different positions – curbed 'contestatory discourse' and in-depth, critical thinking. Erisen et al. (2018), in turn, found in a survey experiment that incongruent information – which can be seen as a proxy for argumentative contestation – increases the integrative complexity of opposing thoughts. Partially corroborating the formal argument made by Chung and Duggan (2020), Baccaro et al. (2016) found that an open discussion format (partly conforming to 'myopic discussion') – while producing the highest amount of opinion change – produced no learning effects, conduced to poorer deliberative quality, and, especially, strong group influence (implying arbitrariness). By contrast, when participants had to take a position (partly conforming to a debate mode), the amount of opinion change was lowest, but there was a significant learning effect, and the deliberative quality (such as justification rationality) was highest.

However, not only is systematic empirical research sparse on this topic, it is also not clear whether contestatory modes represent the only (or premier) route to epistemic advancement. As Rooney (2010) has proposed, collaborative modes – where participants seriously engage with counter-positions and try to 'repair' them in a 'charitable' way – may represent an equifinal pathway to epistemic advancement. Moreover, contestatory modes may be compromised by a number of internal pathologies, such as exacerbating conflict and being blind to potential complementarities of different viewpoints; this, in turn, may compromise the reasoning quality. Nonetheless, we build on the existing theoretical and empirical research and expect that contestatory modes producing the best reasoning quality, defined in terms of argumentative depth, followed by the open and the collaborative mode (which we expect to produce the lowest reasoning quality).

With regard to communication modes and *constructive and integrative thinking,* we expect that collaborative modes trump contestatory and open ones. If one is asked to think through the strengths of other positions, then people frequently put energy into identifying positive elements of these other positions (see Curato et al. 2013). In their laboratory experiment, Schweiger et al. (1986) also identified a trade-off between social acceptability and epistemic advancement. They found that subjects in the consensus groups expressed a higher level of acceptance of their groups' decisions as well as a desire to continue to work with their groups compared to participants in dialectical inquiry or devil's advocacy groups. Overall, we expect the collaborative mode to facilitate the formulation of constructive proposals that bridge different positions and perspectives.

Regarding *opinion change and (de-)polarization*, let us first note that the value of opinion change is contested in deliberative theory (Knight and Johnson 2011; Baccaro et al. 2016). Several theorists have argued that the clarification of opinions is an equally valid (perhaps even more important) goal than opinion change. Habermas (2018, 874) puts this into critical perspective: while he argues that '[c]larifying preferences is, of course, the first step in every political discourse', he simultaneously holds that discourses ground the expectation that the parties should *examine* their initial preferences in the course of deliberation and also change them in the light of better reasons.' Put differently, if people *never change their minds*, then the epistemic value of deliberation is undermined (Schwartzberg 2015, 197; see also De Zúñiga et al. 2023, 328), who state that 'democracy could hardly thrive without changes of mind'. Notice, second, that both the theoretical literature and the empirical findings are fragmented when it comes to opinion change.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

In a recent study, Blumenau and Lauderdale (2022) show that the persuasiveness of rhetorical elements is highly dependent on the context and that there is no single strategy that is most persuasive across multiple policy fields. Indeed, a persistent finding across the different disciplines is that opinion change is highly contingent, dependent on a multitude of mechanisms and highly sensitive to contextual conditions. This combines with a high degree of theoretical fragmentation (for an integrative approach, see Druckman 2022; see also Westwood 2015 for deliberative opinion change).

When we link opinion transformation and (de-)polarization to the three communication modes, there is an expectation in the literature that contestatory modes stifle opinion change (see Groenendyk and Krupnikov 2021). As mentioned above, confronting participants with counter-arguments or counter-positions, contestatory modes might activate 'motivated reasoning' (Taber and Lodge 2006), leading to a 'disconfirmation bias' where participants invest considerable cognitive effort to find arguments that help dismiss the counterargument or the counter-position. Second, being directly challenged with counterarguments can also evoke negative emotions, and people may feel 'defeated and demeaned' (Schweiger et al. 1986, 54). Experimental research shows that outgroup arguments are negatively evaluated when the other side is seen as an opponent or proponent of a contested issue, reducing the listening capacities of discourse participants and consequently stifling open-mindedness. Only when respect in the form of collaborative moves towards other interlocutors was introduced, was this negative effect no longer present (Eschert and Simon 2019). Both disconfirmation bias and negative emotions may produce opinion stability or even opinion polarization. However, some have argued that contestation may have salutary effects for opinion change (Manin 2005, 2017; Deitelhoff 2018): if one is directly confronted with counterarguments, one may not only devote energy to defend and justify own's own position better, but one may also detect inconsistencies in one's own reasoning (see Mercier and Landemore 2012). Moreover, Vargiu et al. (2024) find in an experiment that incivility in communication – whereby incivility is a frequent 'companion' of contestatory modes producing negative emotions on the part of the addressees – does not undermine the persuasiveness of messages. Consequently, it is theoretically and empirically possible that confrontation may also lead to opinion change and depolarization (and does not only produce opinion stability or polarization).

Collaborative modes ask participants to view diverse positions and perspectives – such as health and freedom in our case – as potentially complementary and to recognize possible strengths of diverse positions. This may lead to opinion change and depolarization. Indeed, in a process of citizen engagement carried out in Taiwan in 2019 on the highly controversial issue of regulating Uber cabs, the constructive dialogue resulted in 'recommendations [. . .] that received almost universal approval' ('vTaiwan').[1] Finally, the open mode neither explicitly focuses on debate nor is geared towards finding common ground. We know from empirical research on deliberation that open communication modes – as implemented in deliberative polling (Fishkin 2025) – drive opinion change quite massively, even though it is an open question whether opinion change is driven by the open mode itself or by other factors. Nonetheless, we expect the open mode to produce a higher amount of opinion change than the contestatory mode; however, given that no specific directional trigger is built in, opinions might both depolarize or polarize.

We also expect the three communication modes – especially the collaborative and open ones – to produce higher levels of opinion change compared to the information-only and especially the placebo treatment (with no information and no communication; Esterling et al. 2011; O'Malley et al. 2020). This is in line with findings of deliberative experiments (see Esterling 2018 for an overview), showing that engaging more deeply with arguments drives opinion change. Notice, however, that communication research has shown that opinion change can also happen via more 'peripheral' pathways, such as mere information provision (see Petty and Cacioppo 1986).

[1]For further information see https://www.centreforpublicimpact.org/case-study/building-consensus-compromise-uber-taiwan; https://participedia.net/method/vtaiwan.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

Table 1. Expectations

Dependent Variables	
In-depth Reasoning	A contestatory mode leads to more in-depth justifications compared to a collaborative and open mode.
Integration of Conflictive Perspectives	A collaborative mode facilitates the formulation of constructive proposals compared to a contestatory and open mode.
Opinion formation (Opinion Stability, Polarization, Depolarization)	A contestatory mode leads to more opinion stability and polarization compared to a collaborative and open mode. A collaborative mode leads to more opinion change and depolarization compared to a contestatory and open mode. All communication modes lead to more opinion change than the information-only and the placebo treatment.
Additional factors	
Societal Context	Under conditions of high societal polarization, there is less opinion change and constructivity.
Individual position within the societal debate:	For individuals in the pro-freedom and minority groups, there is less opinion change.

We formulate two further expectations. First, the *societal context* may matter, especially for constructive thinking and opinion change. In a seminal study, Druckman et al. (2013) have demonstrated that the degree of partisan polarization strongly affects opinion formation. Under strong partisan polarization citizens follow partisan cues and ignore arguments that they otherwise consider to be 'strong' (p. 75); in less polarized conditions, they are open to 'strong frames'. Consequently, we expect that under less polarized conditions – as we find them in the first study where the question of prioritizing health v. liberty was contested but opinions were not fully polarized (see below) – the three communication modes affect opinion formation in the way as described before. By contrast, under conditions of high polarization – as in the case of Study 2 on vaccination duty in Austria (see below) – we expect that people's constructivity and open-mindedness are seriously constricted.

Second, we also zoom in on individual heterogeneity. We argue that one's position within the societal debate may affect the transformative potential of participants, especially when debates are polarized (as in the COVID-19 pandemic) and group affiliations and partisan divides get accentuated (Green et al. 2020; Hart et al. 2020; Druckman et al. 2021). Throughout the pandemic, citizens with a pro-health attitude were consistently in the majority, while opponents of governmental infection control measures were consistently in a minority (Wagner et al. 2020; Willems et al. 2020). Psychological 'reactance theory' has been used to shed light on the fact that infection control measures, such as wearing a mask, were perceived as an attack on personal autonomy and freedom of choice (see Sprengholz et al. 2021; Ball and Wozniak 2022), triggering a psychological response that leads the group to seek the restoration of their freedom (Rains 2013). This response might involve discrediting the source of arguments for the necessity of health protection measures, adopting a contrary stance to the message (Rains 2013). In a study by Taylor and Asmundson (2021), individuals who opposed health protection measures were more likely to exhibit psychological reactance. Therefore, we expect that individuals in the pro-freedom minority camp to be less receptive to counterarguments and less inclined to change their minds, especially when overall societal polarization is more pronounced (as observed in the context of mandatory vaccination in Austria in Study 2). We also expect them to be less willing to make constructive proposals to bridge health and freedom positions. Table 1 displays an overview of the expectations for the various communication modes.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

Table 2. Experimental Setup

	Contestatory	Collaborative	Open Communication	Information only	Control
Pre-Survey	Pre-Questionnaire + Positioning for the protection of civil liberties or health				
Information	Pro/Contra Arguments (Freedom v. Health)				Placebo text (general Information)
Communicative intervention	The participants were confronted with a counter-argument to their own position by the AM.	Participants are encouraged by the AM to look for common ground between the two positions.	The participants were asked by the AM for their opinion on the topic.	–	
Treatment Response	Participants wrote short response texts			–	
Post-Survey	Post-Questionnaire + Positioning for the protection of civil liberties or health				

Study 1. Health v. Freedom (Germany)

The first study was conducted in Germany in March 2021, a year after the COVID-19 pandemic outbreak. At that time, Germany was at the beginning of a third 'wave' of COVID-19 infections. However, the population was increasingly protected against severe disease progression by the vaccination campaign. The online survey experiment consisted of a pre-questionnaire, balanced information, a communicative intervention, and a post-questionnaire. Participants were invited to take part in the survey by the survey institute Bilendi: 2,473 people participated in the pre-questionnaire, 2,132 respondents finished the pre- and the post-questionnaire and completed the survey. Overall, the participants in the sample are diverse and approximate the population in terms of age, gender, education, and position on the left-right scale.[2]

An overview of the study design is shown in Table 2. The respondents first had to fill in a pre-questionnaire, where they were asked to position themselves on a 10-point scale with no neutral middle category, either in favour of protecting liberties (1) or protecting health (10). A majority of 67.9 per cent were in favour of protecting health (6-10), while 32.1 per cent of the respondents considered the protection of civil liberties to be more important (1-5); 37.3 per cent of the respondents positioned themselves at the four most extreme points of the scale (1 and 2, as well as 9 and 10). Respondents were then randomly assigned to five different treatment groups. Four groups received balanced arguments for both the protection of health and the protection of liberties, while one group served as a pure control group and received only a placebo text about the COVID-19 pandemic in general.[3] Three groups were randomly assigned to a communicative intervention with an artificial moderator (AM) while one group received the information only.[4] The respondents in the 'contestatory mode' were confronted with a counter-argument according to their own position on health v. freedom in the pre-questionnaire and were asked to take a stance. The respondents in the 'collaborative mode' were asked to search for common ground

[2]The sample approximates the population compared to a representative population survey (see Appendix, Table A-1.1). Only the mean age in our sample is slightly lower and the average education slightly higher and the mean age slightly lower. In Appendix G, the calculations in Figs 1 and 2 were also conducted with a weighted dataset (weighted by gender, age groups, and education levels). The results of the estimates are similar, and the core results are corroborated.

[3]For a list of the displayed arguments, see Appendix, Fig. A-1.

[4]A list of the formulations of the communicative interventions can be found in Table A-2 in the Appendix. See also Fig. A-2 for an example of the communicative intervention by the artificial moderator.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

between the two positions. In the 'open communication mode', respondents were simply asked to say something on the topic. We recognize that the minimal treatments do not represent fully-fledged communication in which participants interact in an iterative way (such as in an adversarial debate with sustained confrontational exchanges between debaters). The rationale behind the minimal treatments was to avoid any communicative dynamics that might occur in iterated group discussion (as we find them in standard deliberative experiments). Such dynamics may sometimes violate the stable unit treatment value assumption (SUTVA) and undermine the ability to draw causal inferences (see Esterling 2018). Moreover, while controlled iterated interaction (steered by an artificial facilitator) could be easily implemented in the context of a debate-style mode, it is difficult to ask individual participants to identify common ground or to say something on a topic several times (which would be necessary to keep treatment intensity equivalent). The randomization worked well – the 2,132 cases are distributed equally among the five treatment groups; there are no significant differences between the groups with respect to sociodemographic variables.[5] After receiving balanced information and being exposed to the various treatments, the respondents were again asked to position themselves on the 10-point scale (pro-freedom v. pro-health).[6] The response from the pre-questionnaire was reproduced in the post-questionnaire in order to help participants remember their initial stance.[7]

Effects on Deepening and Bridging Minds

First, we examine the extent to which the different treatments promote deepening and bridging minds. To do so, we look at the statements that the respondents wrote in response to the communicative intervention. Note that only participants in the three communication modes wrote a statement; if we had also asked participants in the two control conditions to write a statement, then this would have been confounded with an 'open communication mode', where participants were asked to write something after reading the arguments for or against health v. liberty or vaccination duty.

We focus on two indicators of deliberative quality, namely justification rationality and constructive politics (Steenbergen et al. 2003; Gerber et al. 2018). Both indicators are derived from the Discourse Quality Index (Steenbergen et al. 2003; DQI), which represents a philosophically grounded measure of deliberative quality. Justification rationality is an indicator of reasoning quality and measures the extent to which respondents provide complete and extended justifications in their responses. The measure enables us to assess the extent to which respondents have thought through and reflected on their positions; that is, whether the treatments contribute to deepening their minds. It distinguishes between four levels of justification rationality: (0) *No justification*; (1) *Inferior justification* where a reason is given but the linkage between reasons and conclusion is tenuous and incomplete; (2) *Qualified justification* where at least a single complete linkage between reasons and conclusions is made; and (3) *Sophisticated justification* where a problem is examined in-depth by providing multiple and well-justified arguments (Steenbergen et al. 2003).[8] The second indicator, constructive politics,

[5]There are no significant differences between the groups in terms of age, education, position on the left-right scale, as well as knowledge about the COVID-19 virus. Minimal differences exist only with respect to pre-treatment position: individuals in the three groups with communicative intervention are on average slightly more pro-health at the beginning of the questionnaire than individuals in the information-only group. This could be due to the fact that persons belonging to the social minority were more likely to quit the questionnaire early if they had been exposed to a treatment with communicative intervention, in contrast to information only.

[6]The specific wording of the question was: To what extent should civil liberties be restricted in order to reduce the risk of infection in a pandemic? Please tell us on a scale of 1-10 which is more important to you.

[7]However, we acknowledge that this method may also introduce an anchoring effect, potentially encouraging respondents to retain their original answer rather than fully re-evaluating their stance. As such, our results represent rather conservative estimates.

[8]Even though the DQI was originally developed for the analysis of complex argumentation in the context of plenary debates, it also represents a valid measure in the context of the rather short statements, as a comparison with a manually coded indicator 'breath of argumentation' in Appendix H shows.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

captures whether respondents make an effort in integrative thinking and build bridges between their position and the other side. In the original DQI, the measurement contained three codes: (0) *positional politics* where respondents maintain their positions without seeking compromise, reconciliation, or consensus; (1) *alternative proposal* where a respondent suggests a proposal which pertains to a different agenda and does not apply to the discussion at hand; and, (2) *mediating proposal* where a respondent presents a proposal that aligns with the agenda under discussion. We use a slightly modified measure and assign the code *constructive proposal* (1) if an (alternative) suggestion is made regarding what could be done to solve the dilemma between the protection of health and freedom and how to connect both sides or at least what actions should be taken to counter the pandemic and how they could be implemented in concrete terms. However, the proposal must be constructive and directed toward the future; it is not sufficient to merely describe what the respondents think was done wrong in the past and what should have been done instead. The code *no constructive proposal* (0) is assigned if no such constructive suggestion is made. Justification rationality and constructive politics were each coded independently by two people.

The agreement of the manual coding is substantial.[9] For those cases where no agreement was reached in the first coding round, a second coding round was conducted, and a final decision was made. Descriptive statistics of both variables are presented below in Fig. A-3 in the Appendix. A complementary analysis (see Fig. A-6 in Appendix D) suggests that both measures – justification rationality and constructive politics – are independent of each other.[10]

Fig. 1 shows the results of regression models focusing on justification rationality and constructive politics as dependent variables. Model A shows the effects for all respondents, whereas Model B were calculated separately for pro-health protection respondents and Model C for those who were in favour of protecting civil liberties. Results show a positive effect of the contestatory and open communication modes compared to the collaborative treatment (reference category) on justification rationality. With regard to constructive politics, the contestatory and open communication treatments each have negative effects compared to the collaborative treatment. The effects are fairly similar in terms of direction, strength, and significance for all respondents. However, we find that respondents who were in favour of protecting freedom – and who at the same time represented a social minority – made less constructive proposals in the contestatory mode. Notice that the results for constructive proposals might appear slightly tautological at first glance: one might surmise that when being asked to find potential complementarities, then respondents will quasi-automatically do so. But the supplementary bivariate analysis in Fig. A-4 and Fig. A-5 in the Appendix shows that the relationship between the different communication modes, constructive proposals, and justification rationality is non-deterministic. Only half of the respondents in the collaborative treatment actually made a constructive proposal, whereas 11 per cent in the open communication and 5 per cent in the contestatory treatment made such a proposal.

Finally, even though the experimental setup with a large sample does not require the addition of control variables, we have included some additional predictors in the models to check for the validity of our results.[11] Indeed, the control variables corroborate standard expectations (such as a positive relationship between knowledge about COVID-19 and justification rationality).

[9]See Appendix, Table A-3.

[10]Notice that 'integrative complexity' is frequently used as a measure of belief sophistication and argumentation quality (see Brundidge et al. 2014; Erisen et al. 2018). We calculated automated integrative complexity (AutoIC) for respondents' open-ended answers and employed it as a dependent variable in addition to justification rationality (Conway et al. 2014). However, in contrast to our manual codings, no significant effects were found for the treatments (see Appendix, Fig. A-9). We suspect that this is due to the fact that integrative complexity is a compound construct and simultaneously focuses on argument depth and integration. While there is a positive association of integrative complexity measures to DQI indicators (especially justification rationality), the correlation is only 0.57 (Beste and Wyss 2014), indicating that they still measure different things.

[11]Descriptive information on the used covariates is provided in the Appendix, see Table A-4.1.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

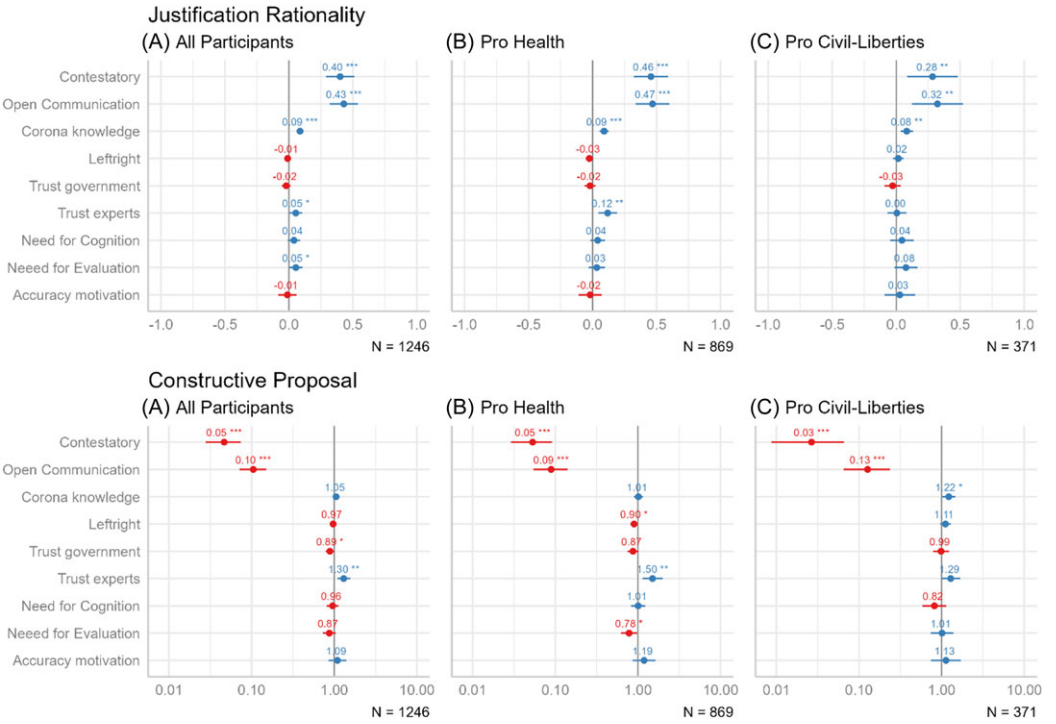

Figure 1. Treatment effects on justification rationality and constructive politics in Study 1.
Note: Linear regression models were estimated for justification rationality and logistic regression models for constructive proposals. Collaborative treatment serves as a reference category; presented are b-coefficients resp. odds ratios.

Effects on Opinion Formation

In order to investigate whether the opinions of the respondents have changed, we compare the pre- and post-survey positions. Overall, there is no aggregate opinion change in either of the two directions; the arithmetic means of the position variable are almost identical in the pre- and post-survey (6.7). This is due to the fact that individual opinion changes cancel each other out and are thus masked (Gastil et al. 2008; Farrar et al. 2010). Hence, we focus on individual opinion changes. The absolute value of the difference between pre- and post-opinions of the respondents is 0.75 scale points, meaning that respondents on average shifted their opinion by 0.75 scale points between the pre- and post-questionnaire (in one of the two directions).[12] Of the respondents, 42.1 per cent shifted their opinion at least by one scale point, and 17.4 per cent shifted their opinion even by at least two scale points. Next, we check how many respondents have moved away from their initial position on health v. freedom (opinion depolarization) or have moved more in the direction of their initial position (opinion polarization). On the basis of the initial position and the post-treatment position, we have calculated a new variable measuring the direction of the opinion change. The new variable has the three values: 'stability', 'polarization', and 'depolarization':

[12]The scale ranges from 1 to 10, with an average change of 0.75 points corresponding to approximately 7.5 per cent of the total scale range. Cohen's d for the absolute change between the pre- and post-questionnaires corresponds to a medium to large effect size.

$$d = \frac{\text{mean}(|X_{\text{post}} - X_{\text{pre}}|)}{\text{SD}(|X_{\text{post}} - X_{\text{pre}}|)} = 0.61$$

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

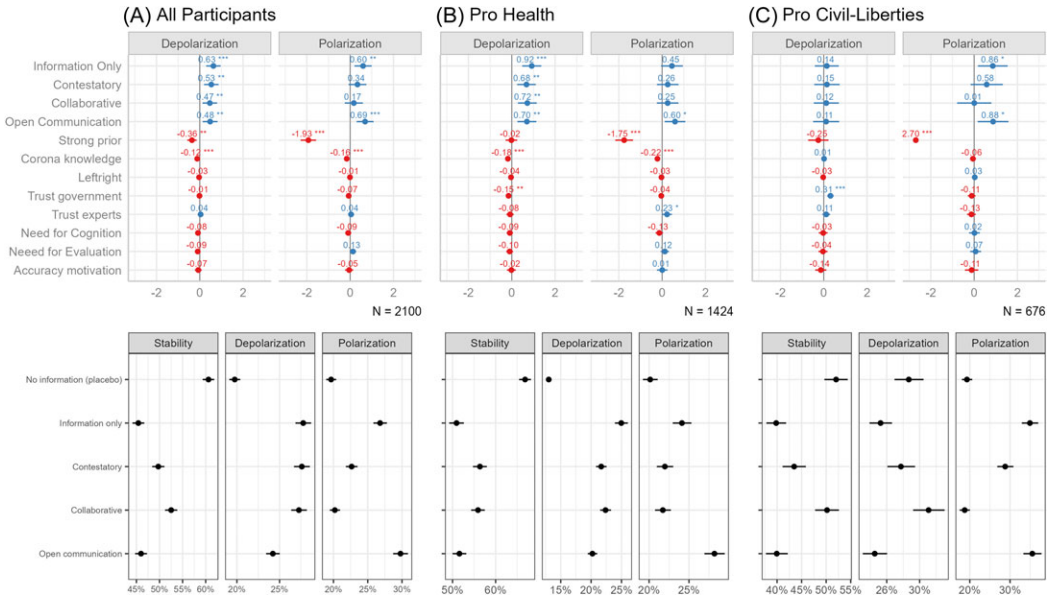

Figure 2. Predictors of Opinion Formation (Study 1).
Notes: Effects on depolarization and polarization compared to opinion stability; the placebo group serves as the reference category for the treatment. The upper panels show the logit coefficients for all independent variables; the lower panels show the predicted probabilities for opinion stability, depolarization, and polarization for each of the five treatment groups.

532 respondents (25.1 per cent) have depolarized their opinions, whereas 362 (17.0 per cent) have become more extreme in their opinions after the treatments. To test whether the different communication modes had an effect on the opinion formation, we estimated multinomial logistic regression models with the three possible outcomes (opinion stability, depolarization or polarization).[13] Again, we added a batch of control variables to check for the validity of our results. These show, for instance, that strong prior attitudes on health or freedom reduce polarization trends (in the form of a 'ceiling effect' where respondents cannot polarize their opinions further on the given scale). Figure 2 displays the results of the regression models. In the upper panels, the logit coefficients of all independent variables included in the models are shown; in the lower panels, the predicted probabilities for the communication treatments are displayed for improved interpretability. We again added some additional covariates to the models.[14] We expect that participants with strong priors, high knowledge, directional motivations, and low trust in government and experts will be both more resistant to opinion change and more likely to polarize the opinions.

In Model A, regression analysis was conducted for all 2100 complete cases.[15] Compared to the placebo conditions, all communication treatments – including the contestatory mode – as well as

[13]One could claim that an opinion change by one scale point is too small to be considered a substantial movement. Nevertheless, we have chosen to measure opinion change as a movement of at least one scale point. There are a few reasons for this: The positioning in the pre-questionnaire was memorized and displayed as 'pre-fill' in the post survey. Thus, the change reflects an actual and conscious movement. Also, respondents changed their position by less than one scale point on average (0.75 in Study 1; 0.3 in Study 2). A one-point change in opinion is therefore already above average. Moreover, only 17.4 per cent of Study 1 respondents and 5.5 per cent of Study 2 respondents changed their opinion by more than one scale point. If we were to define opinion change as movement by more than one scale point between the pre- and post-questionnaire, then the Y-value distribution would be highly unequal, leading to complete separation and non-convergence of the models.

[14]The addition of the covariates has no impact on the effects of the treatments in terms of direction, strength, and significance. Fig. A-10 in the Appendix shows additional models without additional covariates.

[15]2,132 participants completed the pre- and post-survey; however, some variables contain missing values. Regression models were calculated based on all 2,100 complete cases, while some cases were excluded (listwise deletion).

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

the information-only condition conduce to opinion depolarization. Simple engagement with pro- and con- arguments already produces this effect, regardless of any additional communicative intervention. This result tends to contradict findings from deliberative experiments (showing more opinion change in the context of structured discussion compared to mere information; Esterling et al. 2011), but as we discuss below, it partly aligns with findings that good information drives opinion change (Coppock 2022). Regarding opinion polarization, however, a different picture emerges. While the information-only condition and the open communication treatment conduce to more opinion polarization compared to the placebo condition, this is not true for the collaborative and the contestatory modes.

In Models B and C, we separate the analyses for respondents in the pro-health (societal majority; 67.9 per cent in our sample) and the pro-civil liberties camp (societal minority; 32.1 per cent in our sample). Model B shows that for pro-health respondents, an almost identical picture as in Model A emerges regarding the direction, significance and strength of the effects. Results, however, are different when we look at the pro-civil liberty rights respondents in Model C. Here we see that – compared to the placebo condition – none of the treatments has a significant effect on the depolarization of opinions. Opinions within the pro-freedom minority group appear to be already so entrenched that no movement toward the opposite side can be achieved, whatever treatment is assigned. However, the information-only condition and the open communication mode produce opinion polarization for this group.

While opinion changes in the collaborative and open mode conform to our theoretical expectations, opinion changes in the contestatory mode do not (we expected less opinion change and more polarization trends). We found an equal amount of opinion change to the collaborative and open mode, as well as depolarization trends. This may be indicative of a more sophisticated understanding of contestation (Manin 2017; Deitelhoff 2018), claiming that confrontations can lead to opinion revision. Of course, one might object that contestation was deployed in a minimal form (participants were just confronted with a counterargument to their stated position) and that sustained adversarial debating might have produced different effects. However, we have applied another check, namely the number of negative words stated by participants in the open response texts after the treatments. As mentioned above, a standard expectation is that confrontation with counterarguments produces negative emotions (Schweiger et al. 1986; Eschert and Simon 2019). Examining the number of negative words (see Appendix, Table A-5) – controlled for the length of the written statements – we find that the contestatory mode indeed produces more negative words than the collaborative and open format. But despite producing negative reactions, opinion formation in the contestatory mode also leads to depolarization. We will return to this unexpected finding in the discussion section.

Finally, our research design does not allow us to set up a proper mediation analysis testing whether the communication treatments have a causal effect on opinion formation via a deepening or bridging effect. The communicative intervention is a combined treatment consisting of the intervention itself and the response text. Therefore, we can only measure justification rationality and constructive suggestions for respondents in the three treatment groups, but cannot directly compare them with the information only or pure control groups. However, a supplementary bivariate analysis shows that higher justification rationality scores tend to be associated with greater opinion stability (corroborating a previous empirical finding (Baccaro et al. 2016). At the same time, respondents who provided constructive suggestions had a slight depolarization effect (see Appendix, Fig. A-7 and Fig. A-8).

Study 2. Vaccination Duty (Austria)

The second survey experiment was conducted in Austria a few months later, in November and December 2021. The experiment was structured in identical ways as Study 1 in Germany,

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

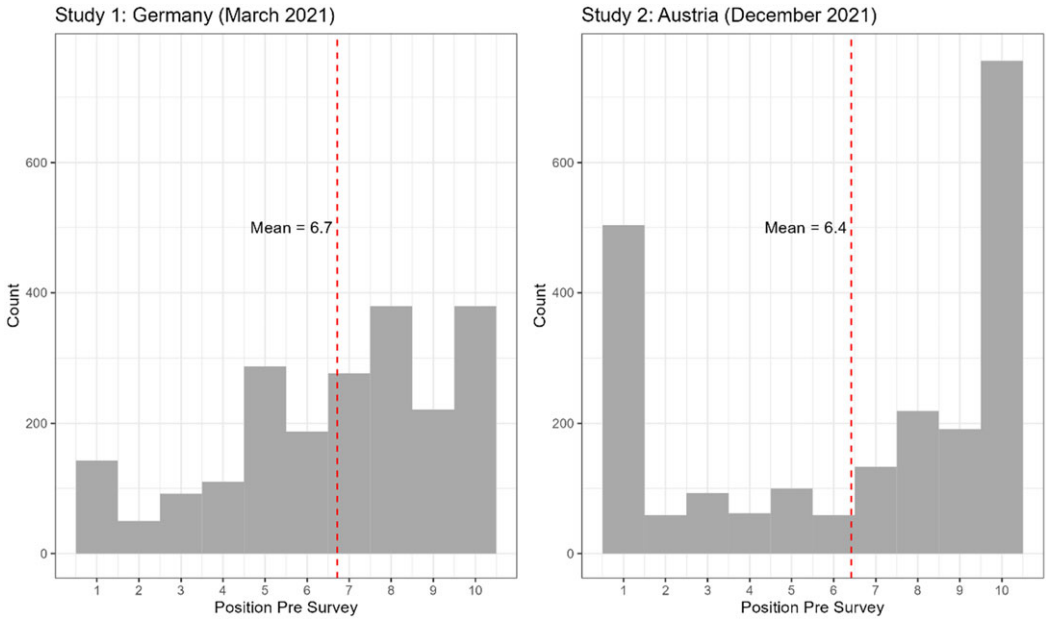

Figure 3. Distribution of pre-survey positions in both surveys.

consisting of a pre-survey, information, a communicative intervention and a post-questionnaire. Again, participants were recruited by the survey institute Bilendi. 2134 people completed the survey.[16]

Compared to Study 1 in Germany, the polarization of the public discourse on COVID-19 had further increased at the time of the survey, and the fourth wave of the pandemic had severely hit Austria. This motivated the government to announce mandatory vaccination, which transformed the rather abstract juxtaposition of freedom v. health into a concrete requirement for individual action, polarizing the debate.[17] In this context, we felt that the original question was too abstract; we also wanted to take advantage of the fact that the introduction of compulsory vaccination is much more polarizing than the issue of health v. freedom, which is why we adapted the question accordingly.

This is reflected in the positioning of respondents on the 10-point scale with no neutral middle category (1 very strongly against mandatory vaccination; 10 very strongly in favour). A large majority of 69.5 per cent of respondents in Study 2 positioned themselves on scale points 1, 2, 9, and 10, and accordingly had a strong prior opinion compared to only 37.3 per cent in Study 1 (see Fig. 3). Despite similar arithmetic means, the two distributions look very different, with only a few respondents in Study 2 positioning themselves in the middle of the scale between values 3 and 8. This distribution creates the very rare opportunity to study the effects of communication modes on deepening, bridging and moving minds under extreme and real-world polarization.

[16]The sample approximates the population compared to a representative population survey (see Appendix, Table A-1.2). Only the mean age in our sample is slightly lower, and the average education slightly higher. Additionally, women are overrepresented in our sample at 59 per cent. In Appendix G, the calculations in Figs 4 and 5 were also conducted with a weighted dataset (weighted by gender, age groups, and education levels). The core results are consistent with the ones presented in Figs 4 and 5 (even though p-values very slightly vary for some of the analyses; for a comparison, see Appendix G).

[17]The legal text is available at: https://www.oesterreich.gv.at/Gesetzliche-Neuerungen/archiv-bgbl-2022/COVID-19-Impfpflichtgesetz-percentE2percent80percent93-COVID-19-IG.html; A documentation of the development is available here: https://viecer.univie.ac.at/corona-blog/corona-blog-beitraege/blog132/; resp. https://viecer.univie.ac.at/corona-blog/corona-blog-beitraege/blog135/.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

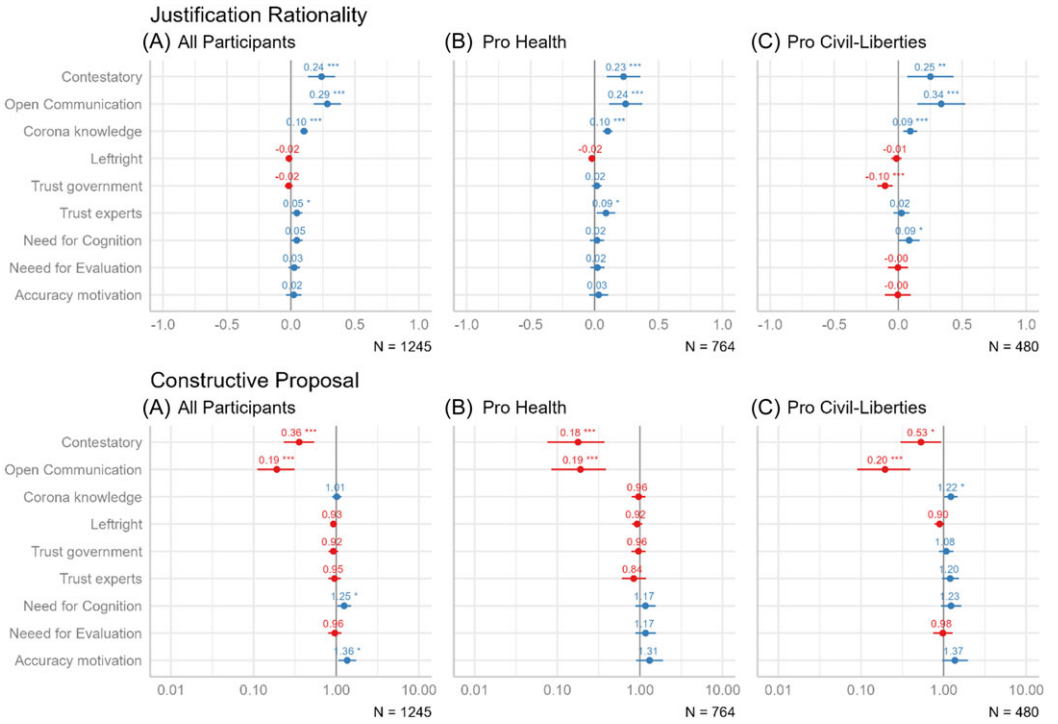

Figure 4. Treatment effects on justification rationality and constructive politics in Study 2.
Note: Linear regression models were estimated for justification rationality and logistic regression models for constructive proposals. Collaborative treatment serves as a reference category; presented are b-coefficients resp. odds ratios.

Finally, both topics – health v. freedom and pro- and anti-vaccination – touch upon the issue of 'protecting health'. As such, we are confident that the comparability of the two topics is basically given. Notice, finally, that the 2,134 respondents who completed the survey were split equally among the five groups; there are also no significant differences between the groups with respect to sociodemographic variables.[18]

Effects on Deepening and Bridging Minds

We first examine the extent to which the treatments have led to deepening and bridging minds. Again, justification rationality, and constructive politics are evaluated on the basis of the respondents' written statements. There was also substantial coding agreement between the two coders.[19] For those cases for which no agreement was reached in the first coding round, a final decision was made in a second coding round.

Fig. 4 shows the results of the regression models with the dependent variables justification rationality and constructive politics. With respect to all respondents (Model A), there is a positive effect of the contestatory and open communication modes on justification rationality, compared with the reference category (collaborative). The effects are almost identical when the pro-health and pro-freedom people are considered separately (Models B and C). As in Study 1, contestatory and open communication have negative effects on the number of constructive proposals. These

[18]Overall, the randomized assignment to the groups worked well: there are no significant differences between the groups in terms of age, education, position on the left-right scale, knowledge about the virus, and pre-treatment position.

[19]See Appendix, Table A-3.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

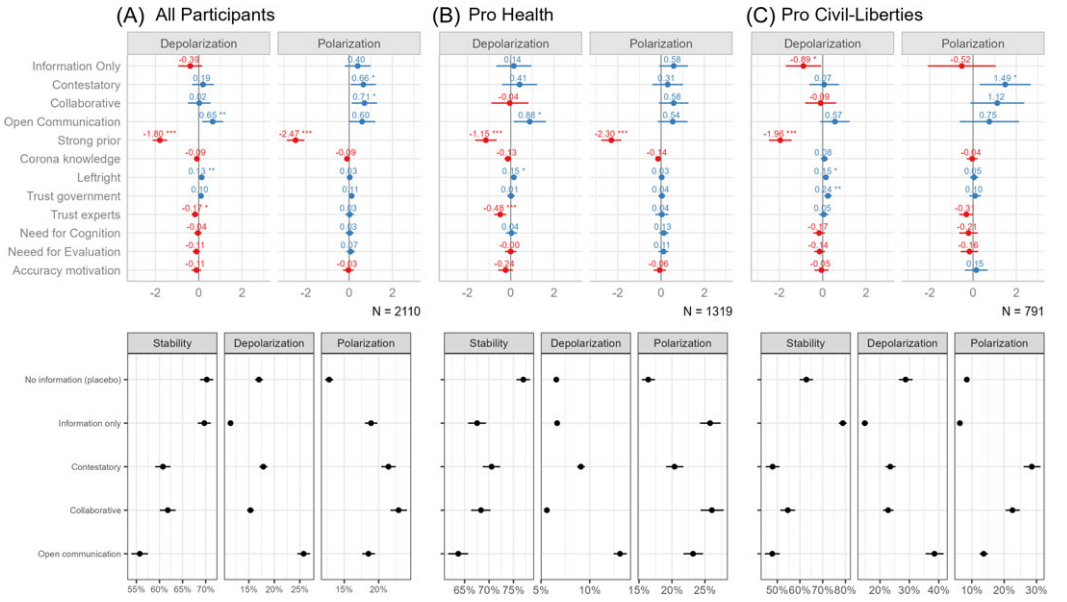

Figure 5. Predictors of Opinion Formation (Study 2).
Notes: Effects on depolarization and polarization compared to opinion stability, the placebo group serves as the reference category for the treatment. The upper panels show the logit coefficients for all independent variables; the lower panels show the predicted probabilities for opinion stability, depolarization, and polarization for each of the five treatment groups.

effects are also comparable for pro-health and pro-freedom respondents in terms of direction, significance and strength.

Effects on Opinion Formation

Subsequently, we examine opinion changes in Study 2. As in Study 1, there is hardly any opinion change at the aggregate level, with an arithmetic mean of 6.4 in the pre-questionnaire and 6.5 in the post-questionnaire. Thus, we again focus only on individual opinion changes. Respondents moved their opinion on average by 0.3 scale points between the pre- and post-questionnaire:[20] 16.2 per cent of the respondents moved their opinion by at least one scale point, and only 5.5 per cent changed their opinion by more than one scale point. Compared to Study 1, these shifts are significantly smaller (in Study 1, the average change was 0.75 scale points). Again, we created a categorial variable to analyze whether participants polarized or depolarized their opinions; 185 (8.7 per cent) respondents depolarized their opinions, 159 (7.5 per cent) respondents polarized, and 1783 (83.8 per cent) respondents did not change their opinions.

Fig. 5 shows the results of three multinomial logistic regressions with a dependent variable differentiating between opinion stability, depolarization or polarization. Displayed are the logit coefficients for depolarization and polarization (compared to opinion stability) in the upper

[20]The scale ranges from 1 to 10, with an average change of 0.3 points corresponding to approximately 3 per cent of the total scale range. Cohen's d for the absolute change between the pre- and post-questionnaires corresponds to a small to medium effect size.

$$d = \frac{\text{mean}(|X_{\text{post}} - X_{\text{pre}}|)}{\text{SD}(|X_{\text{post}} - X_{\text{pre}}|)} = 0.32$$

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

panels and predicted probabilities for the communication treatments in the lower panels. Model A was calculated on all 2110 complete observations.[21]

Unlike in Study 1, information alone does not produce opinion change. In the polarized Austrian case, moderation could only be achieved via the open communication mode. By contrast, both the contestatory and collaborative modes conduce to even stronger opinion polarization (in comparison to the placebo condition). In Models B and C, separate estimates are displayed for those in favour of mandatory vaccination (societal majority; 62.4 per cent in our sample) and those opposed to vaccination duty (societal minority; 37.6 per cent in our sample). Similar to the overall model, Model B shows a positive effect of the open mode on depolarization for supporters of mandatory vaccination. Model C shows that among the opponents of mandatory vaccination, none of the treatment results in significant opinion moderation, whereas the contestatory mode produces polarization tendencies.

In Study 2, participants were also asked about common conspiracy myths about the COVID-19 pandemic. As a robustness check, we ran further regression models with the predictors of opinion formation for respondents with above-average and below-average conspiracy beliefs. As with respondents in the pro-freedom camp (representing the societal minority), the open communication mode also produces depolarization among conspiracy believers (see Appendix, Fig. A-12).

Overall, the effects in Study 2 differ considerably from the ones in Study 1. Open communication is the only mode producing opinion moderation. Intriguingly, the collaborative mode – which we expected to propel opinion change and moderation – produces polarization tendencies instead. The analysis of negative words in the respondents' written texts (see Appendix, Table A-5) again helps to put the results into perspective. In Study 2, both the contestatory and collaborative modes produced more negative words compared to open communication (controlled for the length of the written texts). This may indicate that in the context of high polarization, confronting participants with counter-positions as well as asking them to seek out common ground with the other side is counterproductive and seems to be perceived as a challenge to one' own position. Only a minimal intervention in the form of an open-ended question makes a positive contribution under such polarized conditions.

Again, we included additional control variables in the models (for example, knowledge about COVID-19, prior attitudes), finding that they are in line with standard expectations. To end the section on results, we present an overview of our expectations and findings in Table 3.

Discussion

The results of our study, based on two population-based surveys, present a nuanced and intriguing perspective on the impact of communication modes in the context of varying levels of societal polarization and individual positioning in the societal debate. Surely, while our minimal communicative interventions cannot 'reproduce' how different configurations of 'public debate' (more contestatory or more collaborative) affect the deepening, bridging, and moving of citizens' minds, they nonetheless resemble 'real' and 'designed' *prompts* as we find them in online but also face-to-face discussions: we get confronted with counter-arguments, or we are asked to say something on a topic, or we are asked to find common ground. By testing how these different communication modes affect reasoning and opinion change in a 'pure' fashion, our study cannot only establish causal effects but also shed light on 'deliberation within' in large-scale mass societies

[21]2,134 participants completed the pre- and post-survey; however, some variables contain missing values. Regression models were calculated based on all 2,110 complete cases, while some cases were excluded (listwise deletion).

The addition of the covariates again has no impact on the effects of the treatments in terms of direction, strength and significance. Fig. A-11 in the appendix shows additional models without additional covariates.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

Table 3. Overview of expectations and findings

Hypothesis	Results (Study 1)	Results (Study 2)
A contestatory (and also an open communication) mode leads to more in-depth justifications compared with a collaborative mode.	✓	✓
A collaborative mode facilitates the formulation of constructive proposals (compared to contestatory and open communication).	✓	✓
A contestatory mode leads to opinion stability and polarization	✗	✓
A collaborative mode leads to opinion change and moderation	✓	✗
An open communication mode leads to opinion change (both polarization and moderation)	✓	✓
Communication modes lead to more opinion change than information only	✗	✓
Communication modes lead to more opinion change than the placebo	✓	✓
For individuals in the pro-freedom and minority groups, there is less opinion change.	✓	✓
Under conditions of high polarization, there is less opinion change and constructivity.	✓	

(Goodin 2000) as well as connect with (AI-)interventions designed to improve political communication in the online sphere.

One notable finding is the consistent influence of communication modes on deepening and bridging perspectives, irrespective of the level of societal polarization and individual positioning. In line with theoretical expectations, a contestatory mode, where participants were confronted with a counter-position to their initially stated opinion, had a remarkable capacity to increase depth of argumentation (measured via justification rationality). Similarly, the open mode, where participants were asked to just express their thoughts on the two positions, also contributed to argument depth. This may be surprising given that the formal literature questions the epistemic virtues of ('myopic') discussion (Chung and Duggan 2020). While it may be true that sustained interaction might discriminate more between contestatory and open modes, our results suggest that in a one-shot setup, head-on confrontation is not necessary to generate more in-depth argumentation. The collaborative mode, in turn, where participants were encouraged to collaboratively synthesize divergent perspectives (health and freedom and pro- or anti-vaccination duty), resulted in a greater number of constructive proposals compared to the open and especially the contestatory mode (while simultaneously decreasing the level of justification rationality); this constructive effect of the collaborative mode even materialized for participants in the pro-freedom group (representing a societal minority). This bolsters the finding that small shifts in the presentation of contexts and tasks have quite massive effects on how people think and respond (Groenendyk and Krupnikov 2021), but is also indicative of an important *trade-off* between more in-depth reasoning and more constructive and integrative reasoning.

Our findings regarding opinion change not only challenge traditional understandings of how communication modes exert an influence on public discourse. First, we find that under less polarized conditions, mere information provision already does the job of moving minds. This finding tends to align with Coppock's (2022) approach of 'parallel persuasion' (as well as Goodin's (2000) concept of 'deliberation within'). Parallel persuasion means that people tend to change their minds in the direction of persuasive information, even when that information does not conform to their initial opinions. To be sure, our study design is not in a position to extract or isolate the exact effect of the various pro- and con-arguments presented to participants on their opinion formation: all participants (except the Placebo group) were asked to read both pro- and con-arguments regarding prioritizing health v. freedom and vaccination duty. Consequently, we cannot disentangle the effects of specific (counter-) arguments on opinion change but only assess a 'grand' effect of information. Nonetheless, we note that information *per se* matters for opinion change and perhaps more so than deliberationists tend to assume. Second, under less polarized conditions, we also find opinion depolarization in *all* communication modes, which contradicts traditional expectations claiming that a contestatory mode will undermine moderation trends.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

Again, this is in line with the latest findings in psychology (see Druckman and McGrath 2019). As Benoît and Dubra (2014) write 'when a person is presented with equivocal evidence, that is, evidence that can reasonably be interpreted as being either in favour or against a proposition, [his or her] beliefs can reasonably move either towards or away from accepting the proposition, or not move at all, and by that very fact, the harmonization, moderation, and polarization of two individuals are all reasonable outcomes' (p. 5). However, in line with traditional expectations, the collaborative mode produces less opinion polarization than the contestatory and the open modes. Third, under highly polarized conditions, we find that only open communication produces opinion change. This contradicts traditional expectations that under high polarization, we should deploy collaborative mediation techniques (such as 'appreciative inquiry'; Asif and Klein 2009) to depolarize opinions. Our study shows that such strategies may actually backfire (the amount of negativity is equal to the contestatory mode), which could also be due to the fact that even communication that is not explicitly framed in a confrontative way can be perceived as confrontation (Wright 2022). Only the open questioning has a small effect on depolarization, especially among people in the social majority, but also among conspiracy believers. This result corroborates practical and anecdotal experiences of interacting with (right-wing) extremists: here, communicative 'knock-on'-techniques in the form of open questioning can help create an entry point for more productive conversations.[22] Fourth, in line with previous studies on persuasion, we find that persuasion is contingent and strongly dependent on context (see Blumenau and Lauderdale 2022; Ugarriza and Trujillo-Orrego 2020). The societal level of polarization not only drives the general amount of opinion change, it also structures how different communication modes affect opinion formation. Moreover, individual positioning (pro-health/pro-freedom) in societal debate matters: being in the societal majority (pro-health), people seem to be more open and generous to arguments of the other side; being in the societal minority (pro-freedom), openness is seriously constricted.

Overall, our study contradicts both advocates of a contestatory or a collaborative renewal of public discourse: both have their potential and limits, and neither is in a position to simultaneously deepen, bridge, or move minds. Moreover, an unexpected result is that the open mode – as implemented in many deliberative exercises – tends to perform best *on average* when considering most evaluation standards (argument depth, opinion change, and depolarization), albeit with the exception of constructive thinking. In sum, when it comes to people´s 'internal-reflective' mode (Goodin 2000) and opinion formation, our results point to a 'mosaic' approach to political communication, where the various communication modes are 'activated together in order to enhance the [. . .] multifunctionality of political communication' (Bächtiger 2025). In concrete, this can involve combining and especially *sequencing* contestatory, collaborative, and open communication modes to achieve argument depth, constructivity and opinion change and depolarization.

Conclusion

This study represents the first comprehensive effort to analyse the effects of different communication modes (contestatory, collaborative, and open modes) on argument depth, constructivity, opinion change and opinion (de-)polarization under conditions of higher and lower societal polarization in times of the COVID-19 crisis. We find robust and replicable effects for reasoning quality and constructivity, irrespective of the level of societal polarization and almost irrespective of the individual positioning within this debate: both the contestatory and the open mode produced higher justification rationality (but decreased constructivity), whereas the collaborative mode produced higher levels of constructivity (but decreased justification rationality), indicating a major trade-off between the stimulation of in-depth and constructive

[22]See, for example, Coleman (2011); https://www.mo-asumang.com/mo-trifft-andersdenkende-3sat-sechs-folgen.html.

https://doi.org/10.1017/S0007123425100562 Published online by Cambridge University Press

thinking. In line with previous research, the picture is more complex and fragmented when it comes to opinion change and (de-)polarization. Here, much depends on the level of societal polarization in combination with individual positioning in the societal debate. Overall, there does not seem to exist a simple recipe for moving minds, especially not in stressful times. Future research will need to continue this line of research and especially probe what happens when communication in the various communication modes is *iterated*. In this regard, the various communication modes could also be combined with designed facilitator interventions in deliberative events with citizens. Future research will also need to test our interventions in other societal and political debates, such as migration, climate change or social welfare and check whether our findings generalize beyond the crisis situations such as the COVID-19 pandemic. Given the fact that the COVID-19 crisis affected everyone and given our unique setup with differential levels of societal polarization and people falling into societal majorities and minorities, this study covers a broad range of societal and political situations where our findings might apply. Moreover, our approach with minimal communicative interventions – embedded in more and less polarized environments – sheds light on what John Dewey (1927, 206) considered the central challenge of modern democracies, namely 'the improvement on the methods and conditions of debate, discussion, and persuasion'. We show that the improvement implies *multifaceted* political communications and the recognition of heterogeneous audience needs.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0007123425100562

Data availability statement. Replication data for this article can be found in Harvard Dataverse at: https://doi.org/10.7910/DVN/MMJKNV.

The design of the experiments and some basic expectations regarding the treatment effects were pre-registered and are available at the following links: https://osf.io/kqvjf (Study 1); https://osf.io/a5ysw (Study 2).

Acknowledgements. We thank the 'Optimal Communication'-project team for theoretical, methodological and practical input, namely (in alphabetical order) Seraphine Arnold, Gregor Betz, Elisa Deiss-Helbig, Raphael Heiberger, Felix Kopecky, Jonas Kuhn, Sebastian Padó, Uwe Remer, Anja Rieker, Vanessa Schwaiger and Dominik Wyss (who designed and implemented the artificial moderator). We are also indebted to excellent feedback from the participants at the Yale Conference on 'Deliberative Democracy With a Passion' (May 16, 2024).

Financial support. Terra Incognita Fund of the University of Stuttgart (Project 'Optimale Kommunikation') and financial support from the University of Vienna.

Competing Interests. The authors declare none.

References

Argyle LP, Bail CA, Busby EC, Gubler JR, Howe T, Rytting C, Sorensen T and Wingate D (2023). Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale. *Proceedings of the National Academy of Sciences of the United States of America* **120**(41), e2311627120.

Asif Z and Klein HK (2009). Open and free deliberation: A prerequisite for positive design. *Information and Organization* **19**(3), 186–197.

Baccaro L, Bächtiger A and Deville M (2016). Small Differences that Matter: The Impact of Discussion Modalities on Deliberative Outcomes. *British Journal of Political Science* **46** (3), 551–566.

Bächtiger A (2025). On Perfecting the Conduit of Political Communication: The Virtues and Vices of Contestatory and Collaborative Communication Modes. In Moore A and Bagg S (eds.) *Democracy and Competition. Rethinking the forms, purposes, and values of competition in democracy*. Liverpool: Liverpool University Press, forthcoming.

Bächtiger A and Parkinson J (2019). *Mapping and Measuring Deliberation: Towards a New Deliberative Quality*. Oxford: Oxford University Press.

Ball H and Wozniak TR (2022). Why Do Some Americans Resist COVID-19 Prevention Behavior? An Analysis of Issue Importance, Message Fatigue, and Reactance Regarding COVID-19 Messaging. *Health Communication* **37** (14), 1812–1819.

Benoît J-P and Dubra J (2014). A Theory of Rational Attitude Polarization. *Munich Personal RePEc Archive*, **71648**, 1–42.

Beste S and Wyss D (2014). Cognitive complexity as a proxy for high-quality deliberation? A theoretical and empirical exploration of cognitive complexity and deliberative quality in the EuroPolis discussions: Paper Presentation at the European Consortium for Political Research General Conference, Glasgow, United Kingdom.

Blumenau J and Lauderdale BE (2022). The Variable Persuasiveness of Political Rhetoric. *American Journal of Political Science*.

Brundidge J, Reid SA, Choi S and Muddiman A (2014). The 'Deliberative Digital Divide': Opinion Leadership and Integrative Complexity in the U.S. Political Blogosphere. *Political Psychology* **35**(6), 741–755.

Carcasson M, Black LW and Sink ES (2010). Communication Studies and Deliberative Democracy: Current Contributions and Future Possibilities. *Journal of Deliberative Democracy* **6** (1).

Chen K, Shao A, Burapacheep J and Li Y (2024). Conversational AI and equity through assessing GPT-3's communication with diverse social groups on contentious topics. *Scientific Reports* 14(1), 1561.

Chung H and Duggan J (2020). A Formal Theory of Democratic Deliberation. *American Political Science Review* **114**(1), 14–35.

Coleman PT (2011). *The five per cent: Finding solutions to seemingly impossible conflicts*. New York: PublicAffairs.

Conway LG et al. (2014). Automated Integrative Complexity. *Political Psychology* **35**(5), 603–624.

Coppock A (2022). *Persuasion in Parallel: How Information Changes Minds about Politics*. Chicago: University of Chicago Press.

Curato N, Niemeyer S and Dryzek JS (2013). Appreciative and contestatory inquiry in deliberative forums: Can group hugs be dangerous? *Critical Policy Studies* **7**(1), 1–17.

De Zúñiga HG, Marné HM and Carty E (2023). Abating Dissonant Public Spheres: Exploring the Effects of Affective, Ideological and Perceived Societal Political Polarization on Social Media Political Persuasion. *Political Communication* 40(3), 327–345.

Deitelhoff N (2018). Populismus und Wissenschaft. Streiten gegen das Erlahmen öffentlicher Auseinandersetzung. In Blamberger G, Freimuth A and Strohschneider P (eds), *Vom Umgang mit Fakten*: Antworten aus Natur-, Sozial- und Geisteswissenschaften. Paderborn: Wilhelm Fink, pp. 23–32.

Dewey J (1927). *The Public and Its Problems*. New York: Henry Holt.

Druckman JN et al. (2021). How Affective Polarization Shapes Americans' Political Beliefs: A Study of Response to the COVID-19 Pandemic. *Journal of Experimental Political Science* **8** (3), 223–234.

Druckman JN (2022). A Framework for the Study of Persuasion. *Annual Review of Political Science* **25**(1), 65–88.

Druckman JN and McGrath MC (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change* **9**(2), 111–119.

Druckman JN, Peterson E and Slothuus R (2013). How Elite Partisan Polarization Affects Public Opinion Formation. *American Political Science Review* **107**(1), 57–79.

Dryzek JS et al. (2019). The crisis of democracy and the science of deliberation. *Science* **363**(6432), 1144–1146.

Erisen C, Redlawsk DP and Erisen E (2018). Complex Thinking as a Result of Incongruent Information Exposure. *American Politics Research* **46**(2), 217–245.

Eschert S and Simon B (2019). Respect and political disagreement: Can intergroup respect reduce the biased evaluation of outgroup arguments? *PLOS ONE* **14**(3), 1–12.

Esterling K (2018). Deliberation and Experimental Design. In Bächtiger A, Dryzek JS, Mansbridge J, Warren M and Bächtiger A (eds), *The Oxford Handbook of Deliberative Democracy*. Oxford University Press, pp. 662–677.

Esterling KM, Neblo MA and Lazer DMJ (2011). Means, Motive, and Opportunity in Becoming Informed about Politics: A Deliberative Field Experiment with Members of Congress and Their Constituents. *Public Opinion Quarterly* **75**(3), 483–503.

Farrar C et al. (2010). Disaggregating Deliberation's Effects: An Experiment within a Deliberative Poll. *British Journal of Political Science* **40** (2), 333–347.

Fisher R, Ury W and Patton B (1991). Getting to Yes: Negotiating Agreement Without Giving In. Boston: Houghton Mifflin.

Fishkin JS, Siu A, Diamond L and Bradburn N (2021). Is Deliberation an Antidote to Extreme Partisan Polarization? Reflections on America in One Room. *American Political Science Review* 115(4), 1464–1481.

Fishkin JS (2025). *Can Deliberation Cure the Ills of Democracy?* Oxford: Oxford University Press.

Gadarian SK, Goodman SW and Pepinsky TB (2021). Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic. *PLOS ONE* **16** (4), 1–13.

Gastil J, Black L and Moscovitz K (2008). Ideology, Attitude Change, and Deliberation in Small Face-to-Face Groups. *Political Communication* **25**(1), 23–46.

Gerber M, Bächtiger A, Shikano S, Reber S, Rohr S (2018). Deliberative Abilities and Influence in a Transnational Deliberative Poll (EuroPolis). *British Journal of Political Science* 48(4), 1093–1118.

Goodin RE (2000). Democratic Deliberation Within. *Philosophy & Public Affairs* **29**(1), 81–109.

Goodin RE and Spiekermann K (2018). *An epistemic theory of democracy*. Oxford: Oxford University Press.

Green J et al. (2020). Elusive consensus: Polarization in elite communication on the COVID-19 pandemic. *Science advances* **6**(28), 1–5.

Groenendyk E and Krupnikov Y (2021). What Motivates Reasoning? A Theory of Goal-Dependent Political Evaluation. *American Journal of Political Science* **65**(1), 180–196.

Grönlund K, Herne K and Setälä M (2015). Does Enclave Deliberation Polarize Opinions? *Political Behavior* **37**(4), 995–1020.

Habermas J (2018). Interview with Jürgen Habermas. In Bächtiger A, Dryzek JS, Mansbridge J and Warren M (eds), *The Oxford Handbook of Deliberative Democracy*. Oxford: Oxford University Press, pp. 871–882.

Hangartner D, Gennaro G, Alasiri S, Bahrich N, Bornhoft A, Boucher J, Demirci BB, Derksen L, Hall A, Jochum M, Munoz MM, Richter M, Vogel F, Wittwer S, Wüthrich F, Gilardi F and Donnay K (2021). Empathy-Based counterspeech can reduce racist hate speech in a social media field experiment. *Proceedings of the National Academy of Sciences of the United States of America* 118(50), e2116310118.

Hart PS, Chinn S and Soroka S (2020). Politicization and Polarization in COVID-19 News Coverage. *Science Communication* **42**(5), 679–697.

Himmelroos S and Christensen HS (2014). Deliberation and Opinion Change: Evidence from a Deliberative Mini-public in Finland. *Scandinavian Political Studies* **37**(1), 41–60.

Kerr J, Panagopoulos C and van der Linden S (2021). Political polarization on COVID-19 pandemic response in the United States. *Personality and Individual Differences* **179**, 1–9.

Knight J and Johnson J (2011). *The Priority of Democracy: Political Consequences of Pragmatism*. Princeton: Princeton University Press.

Levendusky MS and Stecula DA (2021). *We Need to Talk: How Cross-Party Dialogue Reduces Affective Polarization*. Cambridge University Press.

Manin B (2005). *Democratic deliberation: Why we should promote debate rather than discussion*.

Manin B (2017). Political Deliberation & the Adversarial Principle. *Daedalus* **146**(3), 39–50.

Mercier H (2016). The Argumentative Theory: Predictions and Empirical Evidence. *Trends in Cognitive Sciences* **20**(9), 689–700.

Mercier H and Landemore H (2012). Reasoning Is for Arguing: Understanding the Successes and Failures of Deliberation. *Political Psychology* **33**(2), 243–258.

Muradova L and Arcenaux K (2022). Reflective political reasoning: Political disagreement and empathy. *European Journal of Political Research* **61**(3), 740–761.

O'Malley E, Farrell DM and Suiter J (2020). Does talking matter? A quasi-experiment assessing the impact of deliberation and information on opinion change. *International Political Science Review* **41**(3), 321–334.

Petty RE and Cacioppo JR (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer New York.

Rains SA (2013). The Nature of Psychological Reactance Revisited: A Meta-Analytic Review. *Human Communication Research* **39**(1), 47–73.

Rooney P (2010). Philosophy, Adversarial Argumentation, and Embattled Reason. *Informal Logic* **30**(3), 203–235.

Schwartzberg M (2015). Epistemic Democracy and Its Challenges. *Annual Review of Political Science* **18**(1), 187–203.

Schweiger DM, Sandberg WR and Ragan JW (1986). Group Approaches for Improving Strategic Decision Making: A Comparative Analysis of Dialectical Inquiry, Devil's Advocacy, and Consensus. *Academy of Management Journal* **29**(1), 51–71.

Sprengholz P, Betsch C and Böhm R (2021). Reactance revisited: Consequences of mandatory and scarce vaccination in the case of COVID-19. *Applied Psychology: Health and Well-Being* **13**(4), 986–995.

Steenbergen MR, Bächtiger A, Spörndli M and Steiner J (2003). Measuring Political Deliberation: A Discourse Quality Index. *Comparative European Politics* 1(1), 21–48.

Stocker S, Bächtiger A, Kittel B and Steenbergen M (2025). Replication Data for Deepening, Bridging, and Moving Minds in Stressful Times. https://doi.org/10.7910/DVN/MMJKNV, Harvard Dataverse, V1.

Taber CS and Lodge M (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* **50**(3), 755–769.

Tannen D (1998). The Relativity of Linguistic Strategies: Rethinking Power and Solidarity in Gender and Dominance. In Cameron D (ed.), *The feminist critique of language: A reader*. London: Routledge, pp. 261–279.

Taylor S and Asmundson GJG (2021). Negative attitudes about facemasks during the COVID-19 pandemic: The dual importance of perceived ineffectiveness and psychological reactance. *PLOS ONE* **16**(2), e0246317.

Ugarriza JE and Trujillo-Orrego N (2020). The ironic effect of deliberation: what we can (and cannot) expect in deeply divided societies. *Acta Politica* **55**(2), 221–241.

Vargiu C, Nai A and Valli C (2024). Uncivil yet persuasive? Testing the persuasiveness of political incivility and the moderating role of populist attitudes and personality traits. *Political Psychology*.

Wagner GG, Kühne S and Siegel NA (2020). Akzeptanz der einschränkenden Corona-Maßnahmen bleibt trotz Lockerungen hoch, *DIW aktuell*, No. 35, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

Westwood SJ (2015). The Role of Persuasion in Deliberative Opinion Change. *Political Communication* **32** (4), 509–528.

Willems J et al. (2020). COVID-19 Pandemie in Österreich – Zufriedenheit der Bürger*innen mit Maßnahmen und Kommunikation.

Wright G (2022). Persuasion or co-creation? Social identity threat and the mechanisms of deliberative transformation. *Journal of Deliberative Democracy* **18**(2), 24–34.

Cite this article: Stocker S, Bächtiger A, Kittel B, and Steenbergen M (2025). Deepening, Bridging, and Moving Minds in Stressful Times. *British Journal of Political Science* 1–22. https://doi.org/10.1017/S0007123425100562