

## Reason against the machine? Future directions for mass online deliberation

Shortall, R.M.; Itten, Anatol; Murukannaiah, P.K.; Jonker, C.M.; van der Meer, Michiel

**DOI**

[10.3389/fpos.2022.946589](https://doi.org/10.3389/fpos.2022.946589)

**Publication date**

2022

**Document Version**

Final published version

**Published in**

Frontiers in Political Science

**Citation (APA)**

Shortall, R. M., Itten, A., Murukannaiah, P. K., Jonker, C. M., & van der Meer, M. (2022). Reason against the machine? Future directions for mass online deliberation. *Frontiers in Political Science*, 4, Article 946589. <https://doi.org/10.3389/fpos.2022.946589>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



## OPEN ACCESS

## EDITED BY

Andrea De Angelis,  
University of Lucerne, Switzerland

## REVIEWED BY

Syeda Naushin,  
University of Malaya, Malaysia  
Mark Klein,  
Massachusetts Institute of Technology,  
United States

## \*CORRESPONDENCE

Anatol Itten  
A.V.Itten@tudelft.nl

## SPECIALTY SECTION

This article was submitted to  
Political Participation,  
a section of the journal  
Frontiers in Political Science

RECEIVED 17 May 2022

ACCEPTED 07 September 2022

PUBLISHED 05 October 2022

## CITATION

Shortall R, Itten A, Meer Mvd,  
Murukannaiah P and Jonker C (2022)  
Reason against the machine? Future  
directions for mass online deliberation.  
*Front. Polit. Sci.* 4:946589.  
doi: 10.3389/fpos.2022.946589

## COPYRIGHT

© 2022 Shortall, Itten, Meer,  
Murukannaiah and Jonker. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Reason against the machine? Future directions for mass online deliberation

Ruth Shortall<sup>1</sup>, Anatol Itten<sup>1\*</sup>, Michiel van der Meer<sup>2</sup>,  
Pradeep Murukannaiah<sup>1</sup> and Catholijn Jonker<sup>1</sup>

<sup>1</sup>Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft, Delft, Netherlands,

<sup>2</sup>Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, Netherlands

Designers of online deliberative platforms aim to counter the degrading quality of online debates. Support technologies such as machine learning and natural language processing open avenues for widening the circle of people involved in deliberation, moving from small groups to “crowd” scale. Numerous design features of large-scale online discussion systems allow larger numbers of people to discuss shared problems, enhance critical thinking, and formulate solutions. We review the transdisciplinary literature on the design of digital mass deliberation platforms and examine the commonly featured design aspects (e.g., argumentation support, automated facilitation, and gamification) that attempt to facilitate scaling up. We find that the literature is largely focused on developing technical fixes for scaling up deliberation, but may neglect the more nuanced requirements of high quality deliberation. Furthermore, current design research is carried out with a small, atypical segment of the world’s population, and little research deals with how to facilitate and accommodate different genders or cultures in deliberation, counter pre-existing social inequalities, build motivation and self-efficacy in certain groups, or deal with differences in cognitive abilities and cultural or linguistic differences. We make design and process recommendations to correct this course and suggest avenues for future research.

## KEYWORDS

digital deliberation, design, automated facilitation, argumentation tools, gamification

## Introduction

While social media and other platforms allow mass participation and sharing of political opinion, they limit exposure to opposing views (Kim et al., 2019), have been shown to create echo chambers or filter bubbles (Bozdag and Van Den Hoven, 2015), and are not effective in reducing the spread of conspiracy content (Faddoul et al., 2020). Furthermore, the policies and design characteristics of popular online platforms are frequently blamed for proliferating discriminatory and abusive behavior (Wulczyn et al., 2017), due to the way they shape our interactions (Levy and Barocas, 2017). Platforms like Facebook or Twitter generally respond to offensive material reactively, censoring it (some argue, to the detriment of free expression) only after complaints are received, which is too late to undo psychological harm to the recipients (Ullmann and Tomalin, 2020). These social media platforms are clearly

unsuited to enabling respectful and reasoned discussions around urgent, important, controversial and complex systemic challenges like climate change adaptation or migration (Gürkan et al., 2010), which are outside the realms of routine experience.

Deliberation offers a decision-making approach for such complex policy problems (Dryzek et al., 2019; Wironen et al., 2019). Deliberative exercises, such as mini-publics, are largely held with randomly selected small groups in both online and offline settings, and offer sufficient time to reflect on arguments, pose questions, or collaborate on solutions. In order to promote effective deliberation online, researchers have proposed that discussion platforms be designed in accordance with deliberative ideals, to increase equity and inclusiveness (e.g., Zhang, 2010) and eliminate discriminatory effects of class, race, and gender inequalities (Gutmann and Thompson, 2004). Not only should they promote a greater degree of equality or civility between participants (Gastil and Black, 2007), they should also redistribute power among ordinary citizens (Curato et al., 2019). Deliberative online platforms ideally strive to promote respectful and thoughtful discussion. Their potential to reduce polarization, build civic capacity and produce higher quality opinions (Strandberg and Grönlund, 2012) is much discussed in the literature. Design features of platforms that promote deliberation are shown to be evaluated more favorably by citizens (Christensen, 2021).

Deliberative discussions, in the most general sense, can be held between any group of people. If based on a randomly selected mini-public model, Goodin (2000) argues, they can accurately reflect the views of a larger group had the process been carried out at that scale. They can provide a certain guarantee of “representativeness” with inclusion of minority voices (Curato et al., 2019; Lafont, 2019). The number of participants in deliberative exercises is typically limited to several hundred at most to maintain organizational feasibility. However, Fishkin (2020) argues that those who do not participate in deliberations are likely to be “disengaged and inattentive”, and are not encouraged to think about the complexities of policy issues posing difficult trade-offs. Simply put, the fewer the people deliberating, the easier it will be to uphold high deliberative standards; but this limits high civic engagement. Conversely, the more people deliberating, the more difficult it will be to safeguard high deliberative standards; but this makes it possible to reach a high civic engagement.

However, obtaining high citizen engagement and maintaining high deliberative standards is particularly appealing now that there is a growing demand for citizen participation in urgent and important policy issues such as climate change or the energy transition (Schleifer and Diep, 2019). And even more so, because valuable and vital public ideas are inadequately reflected in current small-group deliberative practices (Yang et al., 2021). Indeed, the distribution of people’s ideas is shown to have a long tail, thus requiring the participation of masses of people to ensure the diversity of ideas

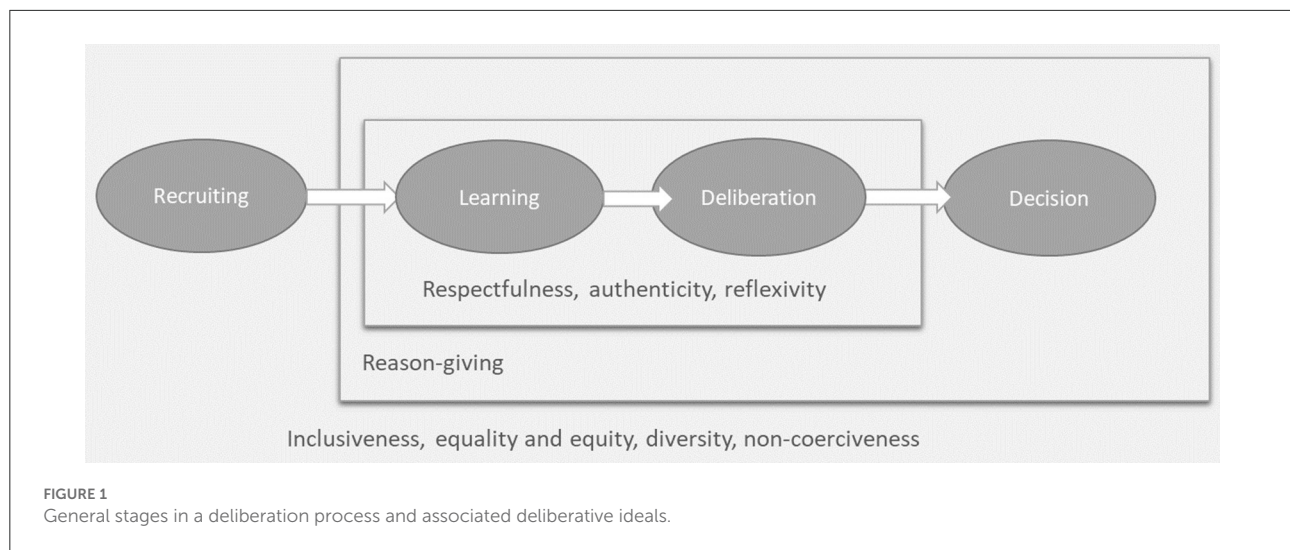
TABLE 1 Summary of challenges to mass online deliberation and deliberative ideals at stake.

Challenges	Deliberative ideals at stake
Lack of participation	Inclusiveness, diversity
Reduced argument quality	Reason-giving, reflexivity, reflection
Empathy loss	Respectfulness, inclusiveness
Power distortions	Equality and equity, non-coerciveness, inclusiveness, authenticity
Negative perceptions	Inclusiveness, diversity, authenticity

is adequately captured (Klein, 2012). There is further empirical evidence that participants from the wider public process the objective information presented in deliberative mini-publics quite differently than the members of this mini-public itself (Suiter et al., 2020). Promoting high-quality mass deliberation online is arguably essential to enhance critical thinking and reflection, to build greater understanding of diverse perspectives and policy issues among participants, while contributing to widely supported solutions (Gürkan et al., 2010; Verdiesen et al., 2018). However, scaling up online deliberation is challenging. In the following, we provide a summary of three main debates in literature relating to challenges of scaling up deliberation to crowd-size, the question about upholding deliberative ideals (e.g., inclusiveness, equity), deliberative tensions (e.g., mass participation vs. quality of participation) and technological solutions to satisfy such ideals. Informed by these debates, we summarize the main challenges to mass online deliberation and their underlying causes in Section Challenges to scaling deliberation online. We then provide a summary of the main ideals of deliberation at stake (Table 1) due to these challenges, which we then use to frame our analysis of the literature.

## Challenges to scaling deliberation online

Deliberative scholars in the past two decades have more or less agreed on a set of ideals that should guide any form of deliberation, such as inclusiveness, respectfulness, giving reason, diversity, equality and equity, authenticity and non-coerciveness (Steenberger et al., 2003; Dryzek, 2010; Mansbridge et al., 2012). More recent work has investigated the extent to which online deliberation fulfills various deliberative ideals, such as civility, respect and heterogeneity, argumentative quality, reflexivity or inclusiveness (Rossini et al., 2021). Figure 1 summarizes the main stages in any deliberative process, along with the ideals of deliberation that should be upheld for each stage. We describe the stages in general terms since deliberative events come in many flavors [for further explanation, see Shortall (2020)], and exact steps may vary. Some ideals can be considered as



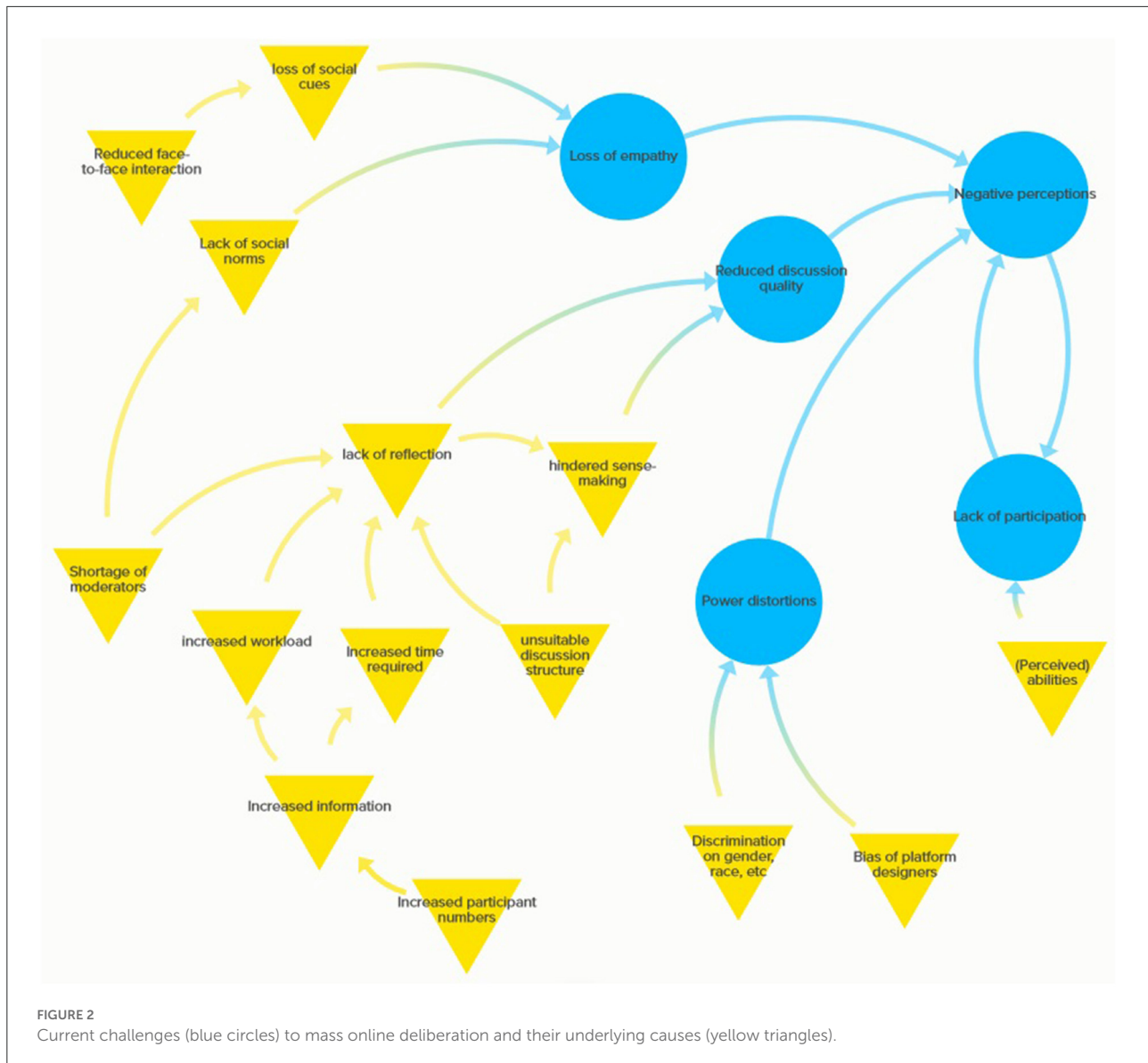
overarching for all stages, whereas others relate mainly to the deliberation stage. In the following, we explain the challenges to mass online deliberation and their underlying causes, which are represented in Figure 2. Table 1 summarizes the challenges and the main deliberative ideals that are at stake for each (We acknowledge that for some challenges, numerous overlapping ideals may be at stake, but for clarity we present only the *primary* ideals, in our view, that are at stake for each challenge).

To begin with, attracting an adequate number of users to deliberate online is not easy (Toots, 2019). Online deliberation spaces may have high barriers to entry for the public (Epstein and Leshed, 2016) and are “typically one-off experiments that occur within the confines of a single issue over a short period of time” (Leighninger, 2012). There are a myriad of reasons why citizens will not take the time and effort to deliberate in depth. For instance, a lack of participation of citizens in online deliberations is related to people’s own abilities, capacities or one’s general perception of online deliberation spaces (Jacquet, 2017). Reduced face-to-face interaction is seen as another component that impedes online deliberation, as it may lead to a loss of social cues and empathy between participants (Rhee and Kim, 2009). A lack of ability to hear, see, or share physical space between participants increases the required effort to communicate and may result in lower mutual understanding (Iandoli et al., 2014). A loss of social cues may also result in a loss of accountability or respectfulness (Sarmiento and Mendonça, 2016). Lack of participation in online deliberation has further been linked to perceptions of deliberation as perpetuating existing power structures (Neblo et al., 2010). In mass deliberation a greater diversity of participants is theoretically possible, however, three main factors of power distortion in online deliberations are said to be gender, social status and knowledge (Monnoyer-Smith and Wojcik, 2012). These factors may not be considered in the design of online

deliberation platforms (Epstein et al., 2014). To the contrary, there is a risk that platform designs may actually strengthen pre-existing (unconscious) biases, i.e., prejudice based on gender, ethnicity, race, class or sexual preference. Other critics note that the perceived legitimacy of deliberative platforms may suffer since they do not uphold certain democratic ideals, such as inclusion, representativeness, equity or equality, which they claim to promote (Alnemr, 2020).

In their review on online deliberation, Friess and Eilders (2015) point out that the design component of online deliberation needs to be better understood, especially how design influences the effects on the individual as well as quality of the deliberation. More recently, Gastil and Broghammer (2021) call for more research on the impact of digital deliberative platforms on institutional legitimacy, which arises from perceptions of procedural justice and trust. Hence, as these scholars show, deliberative ideals are directly influenced by the design features and design process used to create online mass deliberative platforms.

The second debate can be subsumed under the tension of scale vs. quality of deliberation. A first group scholars conclude that the values of deliberation and large-scale participation, and hence the combination thereof - mass deliberation - are inextricably in conflict (Cohen, 2009; Lafont, 2019). Lafont for example notes that open, public accessible deliberation platforms could seriously undermine the quality of deliberation as they would attract attention of populist forces or trolls without a genuine interest into the deliberation itself. Others note that publicity might lead participants to appear tougher and be less cooperative than they really are (Ross and Ward, 1995). Furthermore, since deliberative processes have been designed to encourage slow and careful reasoning, at a large scale they may become too slow to deal with urgent policy problems, such as the ecological crises that we are facing presently, which demand



immediate, urgent action (Wironen et al., 2019). In the current set-up it is understood that the greater the number of people that participate in a deliberation, the less likely it is for everyone to have equal time to explain their views, ask questions and receive answers, weigh new considerations and the like (Lafont, 2019).

A second group of scholars discusses the trade-off between user accessibility and an understandable, well-structured discussion that encourages reflection. Individually, humans are shown to be poor reasoners, although their argumentative capabilities improve when they are encouraged to communicate (Chambers, 2018). Ercan et al. (2019) argue that in a society with abundant information and communication, an important design consideration for online deliberation is to encourage listening and reflection, followed by decision-making. Most online discussions happen on easy-to-use

conversation-based platforms like forums, even though their ability to promote fair and transparent discussion is debatable (Black et al., 2011; Klein, 2012; Fishkin et al., 2018). For example, the structure of comments sections in news media websites is shown to affect the deliberative quality of discussions (Peacock et al., 2019). Another example is the political discourses under a Facebook comment that are of significant low deliberative quality (Fournier-Tombs and Di Marzo Serugendo, 2020). Posts organized temporally, rather than topically are more difficult to navigate and connect to each other and content tends to be repeated. In mass deliberation, without appropriate support, it is impossible for humans to understand and synthesize the large amounts of information that result, let alone reflect on them. Furthermore, disorganized opinions, which may not always be based on facts,

make it difficult to identify and understand the arguments of others.

Hence, it remains not only a challenge to engage the wider population in intensive mass deliberative processes, create meaningful conversations, but also to deal with the fact that those who are expected to deliberate will be unrepresentative of the rest of the population (Fishkin, 2020).

The third broad group of studies focuses on various technical challenges to scaling up deliberation, such as how to manage and facilitate discussion between large numbers (Hartz-Karp and Sullivan, 2014) of users and allow them to make sense of each other's arguments, or how to synthesize the large amount of data that arises (Ercan et al., 2019). Disorganized opinions could be structured by means of argument structuring or reasoning tools (Verdiesen et al., 2018). Visualizing discussions and mapping out arguments helps participants to clarify their thinking and better connect information (Gürkan and Iandoli, 2009; Popa et al., 2020). Such tools or platforms may require user training or supervision, but they also counter sponsored content and promote fair and rational assessment of alternatives (Iandoli et al., 2018). Argumentation tools are therefore a promising avenue for allowing quicker navigation of arguments for participants and encouraging reasoned reflection. Furthermore, having an independent moderator can vastly improve the quality of any discussion, since they can enforce social norms and deliberative ideals. For example, Ito (2018) finds a positive effect on discussion quality due to the 'social presence' of a facilitator on the online platform, along with facilitator support functions.

However, facilitation and moderation is also a challenge when scaling up deliberation. For example, Klein (2012) estimates requiring a human moderator for every 20 users, which theoretically severely limits the ability to scale up to large crowds. Human-annotations and facilitating larger scale online deliberations can be challenging, is prohibitively expensive and resource-intensive (Wulczyn et al., 2017). With mass deliberation, moderator workload becomes too high so some aspects of facilitation may need to be automated to deal with tasks like opinion summarisation or consensus building (Lee et al., 2020). Fishkin et al. (2018) also note a problem with scaling up to be finding recruiting and training neutral (human) moderators. As well as time and location constraints, human moderators also suffer from human bias. Beck et al. (2019), for instance, show how human moderator's beliefs and values may introduce bias into discussions. While support technologies stemming from Artificial Intelligence (AI), such as Natural Language processing (NLP), have been used to aid the moderation and related argument mapping procedures, critics highlight that humans can engage in mutual justification and reflection whereas algorithms cannot (Alnemr, 2020).

Online deliberative spaces remain in the experimental phase and come in all shapes and sizes with no clear consensus about what is the best model. Previous reviews e.g., (Jonsson and Åström, 2014) reviewed research on online deliberation and

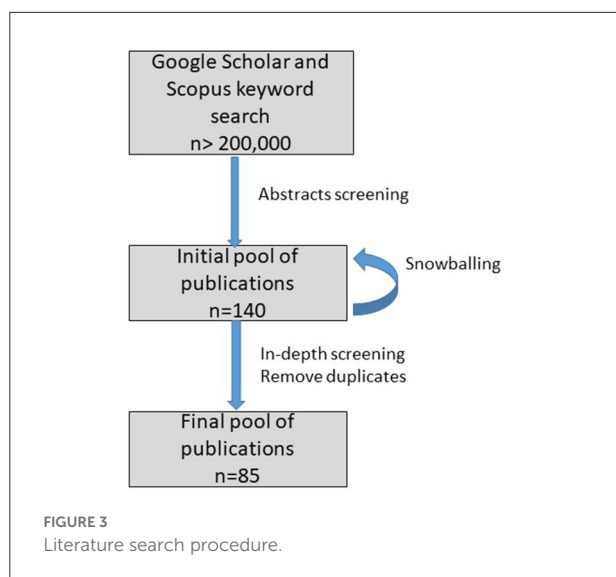
how the concept of deliberation is interpreted by researchers (Friess and Eilders, 2015); analyzed the state of research on online deliberation according to a framework of institutional input, quality of the communication process the expected results of deliberation. While these reviews provide a broad overview of research themes in the online deliberation field, and while design features are discussed to some extent (e.g., mode of communication, anonymity, moderation) neither provide an up-to-date comprehensive systematic overview of existing platform design features. Nor does, to our knowledge, any comprehensive up-to-date review of literature on mass deliberative platforms and their design characteristics exist.

In this review, we therefore explicitly focus on design features of deliberative online platforms that are developed for large groups and aim to analyse the literature with regard to the focus of designers in addressing the aforementioned challenges to scaling up deliberation. We examine the impact of different design approaches on ideals of deliberation, i.e., which ideals are enhanced, and which are possibly detracted from. Our secondary goal is to shed light on the nature of the design process itself, in terms of test users, geographical spread etc. We believe this will provide valuable insights for the future design, legitimacy and hence uptake of online deliberative platforms.

The remainder of this paper is structured as follows: Having described the challenges involved in scaling up deliberation, Section Method presents the method used for the systematic review. In Section Results, we elaborate our findings on the thematic of design features found in the literature according to how they address (or not) challenges of scaling up deliberation. We review the strengths and weaknesses of various design features found in the literature with regard to supporting the scaling up of deliberation online. In this regard, we also discuss the characteristics of case-studies found in the literature, such as the geographical spread and characteristics of test users. Section Summary and future research avenues critically discusses these findings, and offers some points of reflection. Finally, we lay out our conclusions in Section Conclusion.

## Methods

This study aims to provide a comprehensive review of the state of the research of online deliberation platform design. A team of four researchers carried out a literature review on academic (English language) literature relating to digital deliberation platform design. The review consisted of three main steps: 1) literature search, 2) screening, 3) quantitative and thematic analysis. Several searches for academic literature were performed on Google Scholar and Scopus between August 2020 and October 2021. An overview of the method is given in Figure 3. An initial pool of publications was identified using various combinations of the keywords: "digital deliberation",



“online deliberation”, “mass”, “design”. No temporal criteria was applied for the search, given the relative recency of the topic.

Google Scholar and Scopus searches initially produced over 200,000 results, in order of decreasing relevance, which were filtered first by screening the abstracts of publications in the top 20 pages of results. An initial pool of publications was thus identified. This pool was further expanded by snowballing on relevant citations, including using “related literature” links in Google Scholar search. In total 140 publications were found in this stage of the literature search. These were further screened for relevance by in-depth reading in depth by the team of researchers. A database of the selected publications was created and populated with data on the article, journal, design topic or theme, test user numbers and characteristics, country of study, name of platform, main findings, etc.

Relevance screening criteria for the publications in this step were as follows:

- Publication was about an online deliberation platform design case study, i.e., it describes the design or proposed design of a digital deliberation tool or platform, or article analyses existing design(s).
- Publication was peer reviewed or from a high quality source.
- Access was available *via* the author’s institution.

Following this screening step the number of publications was reduced to 85. Publications were then categorized by first clustering literature into design themes. Several major design themes emerged within the pool of publications (Figure 4). The publications were further analyzed to understand how platform designs address (or even worsen) the challenges (see Section Challenges to scaling deliberation online) of mass online

deliberation. Appendix A provides a detailed overview of all publications examined in our review. Appendix B shows a list of commonly mentioned deliberative platforms in the set of publications we reviewed (A list of all reviewed publications is available on request from the first author).

## Results

A total of 85 publications were reviewed (see Appendix A). Of these, 61 were about design case studies and the rest analyzed existing designs or suggested possible design features, e.g., Wright and Street (2007), Towne and Herbsleb (2012), Bozdag and Van Den Hoven (2015), Ruckenstein and Turunen (2020). The most popular design topics were argumentation tools (18 publications), e.g., Iandoli et al. (2014), Gold et al. (2018) and facilitation (16 publications), e.g., Wyss and Beste (2017), Lee et al. (2020). Within the facilitation topic, human facilitation in online deliberation platforms was discussed in 4 publications e.g., Epstein and Leshed (2016); Velikanov (2017).

Gamification was another topic drawing substantial interest with 7 publications, e.g., Gordon et al. (2016), Gastil and Broghammer (2021). Other publications covered the topics of media presentation (3 publications) (Ramsey and Wilson, 2009; Brinker et al., 2015; Semaan et al., 2015) or other specific design dimensions. For example, we found work on virtual reality (Gordon and Manosevitch, 2011), visual cues (Manosevitch et al., 2014) or reflection spaces (Ercan et al., 2019) (1 publication each). The remaining publications (34) focused on general platform design or numerous design features in the same paper. Among these publications, other popular themes of discussion included participant anonymity vs. identity e.g., Rhee and Kim (2009); Rose and Sæb (2010); Gonçalves et al. (2020) and the use of asynchronous vs real-time discussion or text-based vs. video deliberation e.g., Osborne et al. (2018); Kennedy et al. (2020).

Among the publications dealing with specific case-studies, the USA (24 publications) is dominant in this field of research, followed by Europe (23 publications) and Asia (Japan, Korea, Singapore) (8 publications). The remaining publications were from Russia, Australia, Brazil, Israel and Afghanistan (Note: some publications involved research in more than one country).

With regard to the spread across disciplines, around half of the 85 publications were published in what may be considered interdisciplinary, multidisciplinary or transdisciplinary journals or proceedings (e.g., Journal of Information Technology and Politics, Policy and Internet, Government Information Quarterly, New Media and Society, Social Informatics). The remainder were mainly published in either ‘pure’ computer either science (e.g., Applied Intelligence, Journal of Information Science, Journal of Computer-Mediated Communication; Information Sciences) or political science publications (e.g., Journal of Public Deliberation, Policy and Politics, ACM Conferences).

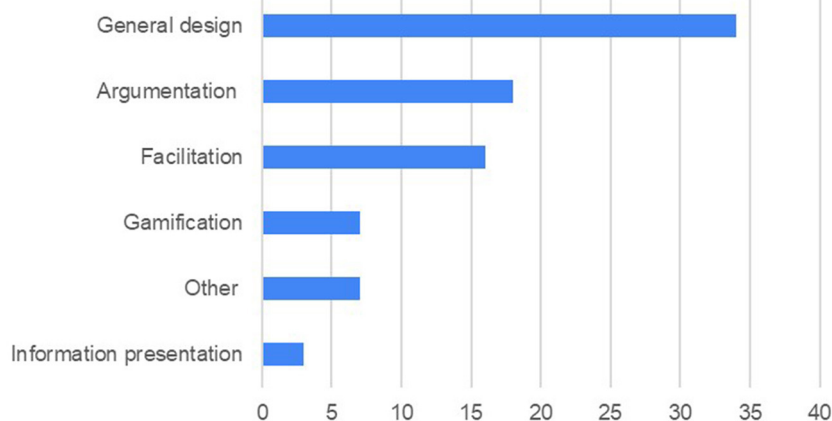


FIGURE 4  
Frequency of design topics.

With regard to the design process itself, it is especially interesting that of the 61 design case studies we analyzed (see [Appendix A](#)), around half had limited or no descriptions of the users or participants who tested or collaborated in the software design. The remaining case studies provided some information on one or more of the following: number of participants or users of the software, (mean) age, gender balance, ethnicity, education levels, political affiliation and profession of participants.

Just 33 case-studies mentioned the number of participants or test users, with numbers ranging from small groups to several thousand users. Only 13 studies provided details about the gender balance of participants. In 9 of these studies, females made up less than 50% of participants. Only 3 studies reported information about the race or ethnicity of participants. The education levels of participants, when reported, was nearly always university level, probably because in the majority of cases where this information was reported, the test users were university students or staff. Only 10 studies reported the age range of participants. While a few studies used samples representative of the general population, for the most part, the age range for participants in these studies skewed young (under 40). In brief, where reported, the majority of test user groups seem to be young, male and WEIRD (Western, Educated, Industrialized, Rich, Democratic).

In the remainder of this section, we present the results of our literature review. We discuss the most common design features found in the literature in relation to how each potentially addresses (or not) the challenges of scaling up deliberation online, as identified in Section Challenges to scaling deliberation online. A summary of our analysis is shown in [Table 2](#), where we explain how each design type potentially promotes or detracts from certain deliberative ideals, with illustrative examples.

## Argument structuring

Aside from publications dealing with general platform design, the highest number of publications we found were about the design theme of argumentation tools for deliberation. Argumentation (where rational dialogue is considered to be the basis of conversation), plays a central role in deliberation ([Fishkin, 2009](#)), and upholding the deliberative ideal of giving reason. During a deliberation, individuals are expected to use reasoning to produce and evaluate arguments from different perspectives ([Mercier and Landemore, 2012](#)). In practical terms, argumentation tools provide a technical solution for speeding up the mass deliberative process by allowing synthesis of large amounts of input when many people are involved in a discussion. These tools also help make it easier to understand the arguments of other participants and may therefore promote a more reflective, reasoned discussion online.

Two common formats of argumentation are argument mapping and issue mapping ([Kunz and Rittel, 1970](#)). Other models include Bipolar Argument Frameworks ([Cayrol and Lagasquie-Schiex, 2005](#)), those using the Argument Interchange Format ([Chesnevar et al., 2006](#)), and various other custom models specific to the use case of the article. So far, no single model is accepted by all. Argument mapping makes the process of producing arguments and their interaction explicit ([Kirschner et al., 2012](#)). Here, the interplay of claims and premises are mapped into a structured format, for instance into a directed graph. However, no single model of argumentation exists, and different theories exist for the structured format ([Reed, 2010](#); [Van Eemeren et al., 2013](#)).

Issue Mapping aims to achieve a similar explicit structure of the content of a debate, but on a more abstract level, in order to create a shared understanding of the problem at hand. Instead of focusing on merely the logical structure of



TABLE 2 Summary of analysis of design types and how they potentially promote or detract from certain deliberative ideals.

Design theme	Ideals promoted	Examples	Ideals potentially adversely affected	Examples
Argument structuring	Reason-giving, reflexivity	Platform provides an easily navigable, logical structure for large numbers of arguments and interactions	Inclusiveness, diversity, equity	Some user groups prefer other forms of expression, or may have difficulty using argument maps
Automated facilitation	Respectfulness, equality and equity, reason-giving	Automated tools help to, e.g., identify abuse, reduce conflict, ask for justifications to arguments or extract discussion structures	Inclusiveness	Moderator bots may lack nuanced skills to engage hesitant participants, or to spot, e.g., gender-based discrimination
Gamification	Respectfulness, inclusivity, diversity, authenticity	<ul style="list-style-type: none"> <li>• Collaborative or interactive designs may increase connection and empathy</li> <li>• Reward mechanisms may increase civility</li> </ul>	Equality, reason-giving, reflexivity, authenticity	<ul style="list-style-type: none"> <li>• Certain designs may inadvertently lead to participants “following the crowd”</li> <li>• Focus on competition may distract from deliberation itself</li> </ul>
Identification	Respectfulness	Identification increases accountability	Equity, diversity, authenticity	<ul style="list-style-type: none"> <li>• Discrimination or harmful social dynamics more likely when social status, gender, etc., are visible.</li> <li>• With identification, some people may hesitate to express their true views</li> </ul>
Anonymity	Equity, authenticity, diversity	Removing visual cues of, e.g., social status or gender may promote equal treatment of participants and hence encourage participation of marginalized groups	Respectfulness	Anonymity may lead to reduced accountability and loss of civility

individual utterances, a debate is mapped out in terms of ideas, positions and arguments (Conklin, 2005) using an Issue-Based Information System (IBIS). While not without its own criticism (Isenmann and Reuter, 1997), IBIS remains a highly popular syntax in recent digital deliberation platforms. This popularity is reflected in the studies included in our search: 8 out of the 18 articles that focus on argumentation design employ IBIS in their platforms.

One of the benefits of traditional social media is that anyone can contribute. Likewise, an easy to browse argument should ideally encourage large numbers of users to contribute (Liddo and Shum, 2013). However, it has been found that argument maps may restrict social interaction between users, due to a significant learning curve. Iandoli et al. (2016) show that these factors have a negative impact on user's experience of platforms with argument maps in terms of mutual understanding, perceived ease of use and perceived effectiveness of collaboration. Similarly, Gürkan et al. (2010) found that users required significant moderation support to input new ideas that fit the formalization of the argument map, and that user activity, such as “normal” conversations, moved outside

of the deliberative platform. Yang et al. (2021) also show that training algorithms to merely define what constitutes a positive, neutral and negative argument (to train automated facilitators that act upon that information), may seriously constrain the space and diversity of opinions. Furthermore, an assumption that unrefuted arguments are winning may unduly influence which arguments are of high quality. Even if no refuting reply is present, it does not necessarily mean that the argument is sound (Boschi et al., 2021). These issues suggest that using argument maps may potentially reduce the overall quality of deliberation unless designs are modified to rectify them.

For example, to preserve social interaction, some platforms mix the argumentation map with traditional conversation-based comments (Fujita et al., 2017; Velikanov, 2017; Gu et al., 2018). To make sure everyone can contribute to the discussion, Velikanov (2017) proposes the use of argumentative coaches. In an attempt to provide further insight into the discussion dynamics, metrics surrounding turn-taking, as well as high level thematic information can be added to these overviews (El-Assady et al., 2016). However, these systems are relatively new, and need more evaluation.

## Automated facilitation tools

Next to the design theme of argumentation, the second highest number of publications we found were about the design of facilitation tools, or automated facilitators. Both design types support online mass deliberation in various ways, by providing technical solutions for large volume tasks, but also by attempting to improve the quality of deliberation by helping people to engage better and managing conflict, for example or by upholding ideals such as “respectfulness”. Automated techniques (e.g., algorithms involving machine learning and natural language processing) can assist moderators with tedious tasks, improving discussions, saving time and give more equal voice to less willing participants. While complete automation, at this point, is still infeasible and will probably impact the quality of the discussion, platforms have been experimenting with such methods, with varying results (Gu et al., 2018).

Studies dealing with human facilitation in online platforms looked at the impact of human moderation on e.g., conflict (Beck et al., 2019), or perceived fairness and legitimacy (Perrault and Zhang, 2019). Conflict may be managed through interface design that reveals moderators beliefs and values before a discussion (Beck et al., 2019). Excessive levels of moderation, however, may reduce perceived fairness and legitimacy because it can lead to self-censorship or the exclusion of under-represented populations. Perrault and Zhang (2019) suggest using crowdsourced moderation to mitigate such negative effects.

Since human moderators may suffer from their own bias or shortcomings, automated facilitation techniques are an important new avenue of research. Epstein and Leshed (2016) found that online human facilitators’ main activities are managing the stream of comments and interacting with comments and commenters. Research on the use of automated facilitation techniques for online deliberation platforms relates to either facilitator assistance tools, which support human facilitators in their roles, or algorithms that completely replace human facilitators.

Tools to assist facilitation tasks may include making sense of large discussions (Zhang, 2017), discussion summaries, recommending contextually appropriate moderator messages (Lee et al., 2020), visualizing dialogue quality indicators (Murray et al., 2013), structuring an issue-based information system (Ito et al., 2020) or analyzing deliberative quality (Fournier-Tombs and Di Marzo Serugendo, 2020). Other tools allow crowd-based idea or argument harvesting (Fujita et al., 2017) or using the machine learning technique, case-based reasoning (CBR), to promote better idea generation, smooth discussion, reduce negative behavior and flaming and provide consensus-oriented guidance (Yang et al., 2019). In that way, the automated facilitation agent extracts semantic discussion structures, generates facilitation messages, and posts them to the discussion system, while the human facilitator can primarily

focus on eliciting consensual decisions from participants (Yang et al., 2021).

Algorithms can also completely replace facilitators and take on the roles of initiating discussions, informing the group, resolving conflict, or playing devil’s advocate (Alnemr, 2020). For instance, (Fishkin et al., 2018) created a platform with an automated moderator “bot” which enforces a speaking queue for participants. The platform also incorporates nudges to encourage participants to follow an agenda. As well as this, active speakers are transcribed in real time and monitored for offensive content. Participants can give the bot feedback about whether to block a user or advance the agenda. Artificial discussion agents can also speed up the learning process for participants on complex issues, help participants to better engage with each other, especially those with opposing views. They might also mine new topics for the discussion from the conversations (Stromer-Galley et al., 2012). For example, Wyss and Beste (2017) designed an artificial facilitator (Sophie) for asynchronous discussion based on argumentative reasoning. They added the automatic facilitator to an asynchronous forum (in an argument tree format) with the purpose of creating and accelerating feedback loops according to the argumentative theory of reasoning. They theorized that this would force participants to recognize flaws in their personal reasoning, motivate them to justify their demands, and ensure that they consider the viewpoints of others. Yang et al. (2021) trained their algorithm on solutions human facilitators used in online discussion systems in the past. This means that the algorithm learns how to identify a given “problem” in a discussion (e.g., “a post or opinion that contains bad words or detouring information that may distract topic posts or main ideas”) and then applies a “solution” (e.g. “remove such a post from discussion or hide it in order to smooth the forum discussion”).

While tools and algorithms are shown to be highly useful and reduce the workload of human facilitators in large-scale deliberations, it is a risk that human bias is replaced with the inherent bias of the algorithms behind automated moderators. Nonetheless, only a few authors in the literature review take up this discussion. Alnemr (2020) critiques the developments of Stromer-Galley et al. (2012) and Wyss and Beste (2017), noting that algorithms underpinning automated facilitation lack the ability to use discretion as humans do, about how to ensure inclusion, and to enforce certain deliberative norms in a sensitive way. She argues that expert-driven design does not allow citizens to define and agree on deliberative ideals and that instead these ideals are imposed on them by the algorithm, whose design is not transparent.

In the process of facilitating, moderators may also impose their own biases into the conversation. While users themselves could be tasked with inputting their ideas into an argument map (Pingree, 2009), the mapping procedure involves multiple cognitively challenging tasks, increasing the risk for errors. Mappers (those doing the mapping) are expected to keep up

with the conversation, while keeping close attention to the map and perform clarification interactions with the discussants. In a review of Finnish moderator tools in various platforms, Ruckenstein and Turunen (2020) warn that some designs may force moderators to operate more like machines and prevent them from using their own skills and judgement, leaving them frustrated in their work. Wyss and Beste (2017) caution that, when designing (automated) facilitators, it is crucial to find a balance between interfering and not interfering in the discussion and this can depend on the context. More research in this regard is clearly needed.

Three publications deal with the topic of learning styles or argumentative capability of participants: Velikanov (2017) makes some preliminary technical and non-technical design suggestions (e.g., argument coaches or facilitator incentives) to address the topic of differing levels of factual preparedness and argumentative capability as well as the issue of linguistic differences. Epstein and Leshed (2016) described heuristics that can be used by moderators to respond to different levels of participatory literacy and hence improve the overall quality of the comments. Brinker et al. (2015) found that using mixed media providing favorable outcomes for developing social ties, building mutual understanding and encouraging reflection on values as well as facts, and formulating arguments. In their study, Wyss and Beste (2017) show that the effectiveness of the automated facilitator in creating favorable learning conditions for participants depends also on their personality traits, e.g., on whether participants were conflict-avoidant or had a high need to evaluate.

Furthermore, few studies deal with the issue of gender, or gendered behavior in relation to facilitation approaches or the design of automated facilitation tools in deliberative software. One study, (Kennedy et al., 2020) found that the gender of human facilitators impacted discussion outcomes, that a text-based environment may be favorable to female participants since it prevents interruptions, but that non-white and participants over 65 were less active in such discussions.

## Gamification

Gamification of deliberation platforms has gained attention in recent years, showing potential to uphold deliberative ideals like civility or to engage more citizens in online deliberation as well as in building empathy. Hassan (2017) hopes that gamification would increase low levels of civic participation, and motivate public officials and government agencies to engage in digital deliberative projects as well as reducing negative sentiments toward such projects. She assumes that if more citizens are encouraged to participate in deliberations and perceive their involvement as meaningful, it would not only increase the number of citizens in the governance of their communities, but also help to convince public officials that it can

improve governmental decision-making (Hassan, 2017). Beyond engagement and activation, gamification is seen to improve the quality of digital deliberations by promoting social connection and empathy, improving self-esteem and social status, argue Gastil and Broghammer (2021). To create empathy for example, the authors suggest that a deliberation bot could ask participants to reflect on the needs and experiences of fictionalized residents of their community or discuss fictional scenarios based on real residents.

The functionalities of gamified environments for deliberation in the literature are many. Some use reward-based approaches to award users based on their activities, or stimulating inter-user competition (Hassan, 2017) others may use collaborative reward structures that provide incentives for reaching a consensus or taking collective action (Gastil and Broghammer, 2021). Others again try to steer users to formulate opinions that fit the discussion phase and to shorten the response time of answering questions (Ito, 2018), or to promote civility among users (Jhaver et al., 2017).

Popular reward functions that have been applied in digital deliberations are deltas (Jhaver et al., 2017), badges (Bista et al., 2014), award scores (Yang et al., 2021) or discussion points/virtual money (Ito, 2018). Gastil and Broghammer (2021) provide a further comprehensive list of gamification mechanisms suitable for online tools that include deliberation, such as competition over resources; material rewards; scarcity of rewards; artificial time constraints; chats; peer rating system, shared narratives; learning modules; progress bars; leaderboards and missions.

Few case studies have been conducted so far that measure the effects of gamification. Bista et al. (2014) for example used a set of badges as awards for a gamified community for Australian welfare recipients who were encouraged to communicate with each other and with the government. According to the authors, gamification increased participant's reading and commenting actions, and helped the researchers to obtain accelerated feedback on the "mood" in the community. A recent review found however that it is difficult to conclude the impact of gamification in online deliberations, since gamification is never equally experienced and appreciated by all sorts of social groups in the same way (Hassan and Hamari, 2020).

While conducting field experiments for gamification in online deliberation, Johnson et al. (2017) found issues with equality, turn taking, providing evidence in discussions, elaboration of stakeholder opinions and documentation. After inserting a turn-taking mechanism, combined with restricting the topic of conversation through "cards" prevented talks from going off-topic. Moreover, reasoning prompts improved empathy and respect, as well as the change of viewpoints following some discussions. This also reduced the workload for facilitators, as groups of participants regulated themselves according to the rules of the game.

Jhaver et al. (2017), who studied the subreddit ChangeMyView, observed that gamification mechanisms in the form of deltas increase civility and politeness between users. This is particularly obvious for new members of the community as well as for high-scorers. Even though there is a competitive element that allows high performers to compete with one another, the focus of the community is on meaningful conversations, and thus, according to the authors, the users are not explicitly judged by their delta scores. These findings are in line with Hamari et al. (2014) that conclude that positive outcomes from gamification depend on the context and the characteristics of users.

Downsides of gamification methods have also been reported: features that allow participants to build reputation *via* up-voting for instance. While having voting mechanisms may seem democratic, in reality this could be far from the case. Humans tend to be highly influenced by the opinions of others: for instance, studies show that comments with an early up-vote are 25% more highly rated in the end, regardless of their quality (Muchnik et al., 2013; Maia and Rezende, 2016). Some users might also be driven solely to achieve high scores and gaming the gamification rather than to actually contribute to the deliberation goal (Bista et al., 2014). Such occurrences might lead designers to “hide” the full rules of calculating points to participants, which then raises issues of transparency.

Gamification can be an important addition to developing digital deliberation interfaces, and if embedded with social norms, and moderation mechanisms, it can improve a digital space to foster productive discourse. Competition generated by gamification should never be upfront. Nor should it be unidirectionally focused on those members who are already achieving strong results (Jhaver et al., 2017). Civic gamification designers should focus on understanding the psychology (intrinsic motivation) of why citizens would participate in an online deliberation in the first place, for gamification efforts to provide value (Hassan, 2017). Moreover, gamification leads to more data processing, which does not necessarily mean that this data informs decision-makers in ways that represent public opinion. Cherry-picking and biases might occur in terms of privileging what they see as “facts”, and “evidence” such as “user engagement”, “likes” or “scores” etc., because decision-makers succumb commonly more to these types of aggregated data than to qualitative attributes (Johnson et al., 2017).

## Anonymity and identifiability

Several of the remaining publications pay attention to design aspects of anonymity and identifiability and the impact of social cues on social dynamics in online deliberation platforms. Choosing between identification or anonymity in digital deliberations creates a number of trade-offs with regard to upholding deliberative ideals.

Anonymity can imply a loss of accountability and respectfulness or may negatively influence civility in online discussions (Sarmiento and Mendonça, 2016). On the other hand, with anonymity, a more egalitarian environment is possible since people feel more freedom to express their honest, even if unpopular, point of view. Removing visual cues of gender, age and race promotes equal treatment of individuals (Kennedy et al., 2020). Harmful social dynamics can be reduced and people stay more focused on the task at hand (Iandoli et al., 2014). Anonymity can also allow civil servants or people with neutrality obligations to participate. Asynchronous discussion is also a way to “level the playing field” between the more and less informed public (Neblo et al., 2010), and has been shown to encourage women to participate more by removing interruptions (Kennedy et al., 2020).

Reducing anonymity may have a positive effect on respectfulness and thoughtfulness (Coleman and Moss, 2012) and increases transparency, but has a possibly negative effect on engagement (Rhee and Kim, 2009) — people tend to contribute less to the discussion overall when they are identifiable. This can also raise questions of privacy on the platforms. Gonçalves et al. (2020) indicate how easy it is to build a profile of each participant by following, or studying the tracks they unconsciously leave behind while discussing, proposing ideas or interacting with others as well as their behavior in a gamified environment.

In addition, it has also been found that when users share a common social identity in an online community, they are more susceptible to group influence and stereotyping, despite participant anonymity (Jhaver et al., 2017). Incorporating the focus theory of normative conduct (Cialdini et al., 1990) into platform cue design may improve online dialogue by enforcing three different types of norm: descriptive, injunctive and personal. Respectively, these norms motivate behavior by promising either rewards or sanctions externally imposed by others; by providing frequency information about the behavior of others and reflecting individuals’ commitment with their internalized values. Manosevitch et al. (2014) show that visual/cognitive cues like banners can prime participants to be aware that they are in a deliberative context. Such cues may encourage e.g., reflection, considering a range of opinions, and being true to one’s self and hence improve the deliberative quality of the discussion.

## Summary and future research avenues

Six years ago, in Friess and Eilders’ online deliberation review (Friess and Eilders, 2015), the number of mentions about automated facilitation, argumentation mining, algorithms, or gamification was exactly zero. Much has changed since then. As this literature review shows, the development and experimental use of mass deliberative platforms has leapt forward in recent

years, and has not only seen increasing scholarly attention, but also a rise of tech-start-ups. We have found references to 106 different deliberative online platforms in our review (see [Appendix B](#)). Certainly, also due to the COVID-19 pandemic, the world has experienced a massive increase of communication and collaboration online and the results of this will soon come to light. This presumably increased the overall computing literacy and also created new communication and cultural habits that might positively impact the development of mass online deliberation.

As we have focused on designers and their progress in addressing the challenges to scaling up deliberation online, our review highlights first and foremost that automated facilitation and argumentation tools receive by far the most attention in the literature surveyed. Compared to the review of [Friess and Eilders \(2015\)](#), which considered “design” in a broader sense, including also the conceptual design of the deliberation process, we now see an increasingly narrow focus on the design of technologies, in order to deal with the large number of users and information produced.

Since argumentation tends to be regarded as a key goal of deliberation, as a means of promoting deliberative ideals of reflective, reasoned discussion, mining and mapping tools are currently the focus of extensive research and experimentation in the area of argumentation. An interesting discrepancy of this process is that while years of deliberation research has gone to great pains to conceptualize what makes a good argument, or what is crucial for deliberative quality or how deliberation can be applied in a given context or culture, today’s tools do not necessarily reflect this. Current efforts that train an algorithm to detect and label an argument often resort to pre-existing popular syntax, such as Case-Based Reasoning (CBR), Issue-Based Information System (IBIS), Bipolar Argument Framework (BAF) or Argument Interchange Format (AIF). While we have not compared the syntax behind crowd deliberative platforms to deliberative ideals in this research in detail, it would be an avenue of research to investigate if these syntaxes form a new understanding of deliberative ideals and if they are at cross-purposes with older definitions. Moreover, a perspective found to be missing in our literature search relates to a critique of logical argument structuring [e.g., [Durnova et al. \(2016\)](#)], which argues that a discursive approach may be more inclusive appropriate since it recognizes the subjectivity of actors, their different forms of knowledge and interpretations from which they create meaning. It is shown, for example, that rather than sharing facts, people are often more convinced by stories about personal experiences ([Kubin et al., 2021](#)), which suggests that deliberative software that relies on the exchange of arguments, or factual statements may not achieve the other desired goals of deliberation such as building respect for different points of view or empathy. Introducing the possibility for sharing other types of information in a discussion, or incorporating game elements

may improve these aspects and encourage a more diverse range of users to participate.

Automated facilitation tools also receive a lot of attention. As our review shows, they are mostly being used to tackle problems related to scaling up deliberation such as managing comments, maintaining respectfulness, encouraging participation, making sense of vast amounts of information or monitoring the discussion progress and quality. However, while such tools may aim to circumvent human facilitator bias, issues of algorithmic bias are hardly discussed in the publications we found. There is very little discussion about the values or interpretation of deliberative democracy that underpins automated facilitation tools. Moreover, reformulating provocative content, mirroring perspectives, posing circular questions or playing the devil’s advocate, are only a few of the important techniques a trained facilitator can employ to support argumentatively disadvantaged participants or to redirect a discussion that reaches a dead end. In our literature review, we find that automatic facilitation tools currently lack the functionality for such nuanced interventions on a large scale, suggesting that scaling up may result in a loss of deliberative quality. Hence, as our review shows, much work remains to be done in relation to tools that assist with cognitively challenging tasks involved in mapping the arguments, deciding which arguments should win, teaching diverse participants to use argumentation itself and allowing for social interaction in parallel, so that such platforms can be as attractive as popular social media. Some possible solutions include providing opportunities for social interaction alongside argumentation, providing argumentation training for users or coaching for less confident users so that they feel on a par with more articulate or privileged users.

As we see, although facilitation and argumentation tools may overcome certain challenges to mass deliberation, they may neglect others, or create new problems. Therefore, the investigation of other design formats are also important. Substantial attention is now given to gamification features in the literature, which address a different issue, that of encouraging participation online and fostering feelings of connection and empathy. Some research has begun on how to accommodate different communication and learning styles, the use of gamification as an addition to argumentation or a way to build rapport between online participants. Experimental work on different gamification designs (e.g., competitive design, collaborative design, etc.) shows promise, in particular in increasing engagement. There are further indications that gamification can increase engaged interaction when it is designed around contextual and user-specific means, rather than around what is popular in gamification research ([Hassan and Hamari, 2020](#)). Further research on how gamification can promote different communication styles or different deliberative ideals is needed. Until now, literature fails to demonstrate a wide array of comparative studies ([Hassan, 2017](#)) and it is unclear

how different user-groups utilize gamification functions to their advantage in certain contexts.

If we compare our findings against the literature on deliberative ideals, we find that especially themes such as inclusiveness, diversity and equality lack attention in papers of online mass deliberation. Our analysis of the literature (See Section Results) suggests that most research on deliberative platforms uses highly educated test users, likely from a comfortable economic background. The geographical distribution of studies shows that a high proportion of research on the design of deliberative platforms takes place in Western, developed countries, and rarely reports the ethnicity of the participants, which will doubtless have implications for the design of future platforms and should be a concern. In particular, Western cultural and methodological standards currently dominate deliberation research (Min, 2014). However, certain cultures may favor different deliberation or argumentation styles e.g., cultures may be consensus-based or adversarial (Bächtiger and Hangartner, 2010). For example, Confucian inspired societies may value social harmony over public disagreement. Other studies (Becker et al., 2019; Shi et al., 2019) have highlighted the positive impact of partisan and polarized crowds. Moreover, in Muslim countries men and women may deliberate separately (Min, 2014). Elsewhere, in societies where ethnic, religious, or ideological groups have found their own identity in rejecting the identity of the other, online deliberation needs to find common ground first (Dryzek et al., 2019).

The lack of diversity in mass online deliberation studies and test users points to further problems. Internet use is influenced by less than obvious factors such as social class or sense of self-efficacy, which often affect people of color or women (López and Farzan, 2017) or motivation, access to equipment, materials, skills or the social, political and economic context (Epstein et al., 2014). In practice, difficulties may arise with including certain social groups (minorities, the poor or less-educated) in online deliberations (Asenbaum, 2016). Some citizens may believe complex issues are beyond their expertise and hence defer decisions to be made by experts (Font et al., 2015) and people may need training or education to help develop the capabilities to participate in a deliberative setting (Beauvais, 2018). Too heavy a focus on technical solutions and designs may hence exclude people with limited argumentation and rhetorical skills or that prefer other types of expression. Future research should have a stronger focus on how populations with special needs may be better included in online mass deliberation, especially older adults, children, people with disabilities or those who are illiterate (López and Farzan, 2017). Language barriers may also be an issue of further research, as in that if not all participants speak the same first language, this may lead to further inequalities (Velikanov, 2017).

Furthermore, little attention is paid in the literature surveyed to the effects of gendered behavior (Afsahi, 2021),

social inequality (Beauvais, 2018) or the implications of different communication or learning styles (Siu, 2017), all of which can influence inclusiveness and equality in deliberation. Communication style and gendered behaviors are transferred over into deliberative settings. The communication skills and style of expression required in deliberative settings tends to be characteristic of higher income white males and may not be characteristic of all social groups, leaving them at a disadvantage. For example, high-resource and digitally engaged individuals are generally more capable and active in discussions (Himmelroos et al., 2017; Sandover et al., 2020), leaving less-privileged, less-educated, or perhaps illiterate participants at a disadvantage in discussions with the more privileged, better educated, and articulate (Hendriks, 2016; Siu, 2017). Furthermore, men are more likely to interrupt or ignore women and speak for longer and more often than women, who tend to be more passive or accommodating in deliberations (Siu, 2017). Women's contributions may be marginalized or sexualised due to gender hostility (Kennedy et al., 2020). Women also tend to be more conflict-avoidant and less willing to engage in argumentation required for deliberation, hence they may require a particular style of facilitation to ensure their inclusion (Afsahi, 2021). Our analysis of the literature suggests that women tend to be underrepresented in the groups of users that test deliberative platforms during the design process. This is a concern and should be considered in future design studies.

Lastly, as our literature review shows, the tension between scale vs. quality of deliberation is frequently being circumvented by narrowing the focus of deliberative quality, for example highlighting metrics such as the amount of contributions, continuous user interactions (Ito, 2018), reaching a consensus or taking collective action (Gastil and Broghammer, 2021), the absence of flaming or irrelevant content (Yang et al., 2021) as well as increased civility and politeness between users (Jhaver et al., 2017). Obviously, by scaling up online deliberation, designers focus on one problem at a time by testing different design interventions on their effectiveness, hence the quality of mass online deliberation is likely to increase over time. Yet, as mentioned earlier, the one-sided use of highly skilled test-users may hide problems related to the quality of deliberation, such as trolling, spreading of disinformation or increased polarization that might only occur if mass deliberation software is applied on a large-scale and used across different groups of society.

As pointed out by Friess and Eilders (2015), research about how to reconcile individual preconditions, such as opinion change or mutual trust with quality and legitimacy on the collective level are still largely missing in crowd size deliberations, and should be a source for further research. However, the amount of generated data by mass deliberative platforms is already estimated to be huge. It is thus likely that in a short period, such research will be very likely to follow. Caution has to be exercised, however, when participants' behavior is surveyed, paternalised or nudged into a certain direction, to

achieve what Hassan and Hamari (2020) call, the “creation of good citizens”. Protocols for how we speak and interact with each other online may be increasingly guided and determined by algorithms, whose design is not often transparent (Alnemr, 2020).

## Conclusion

This paper reviewed the literature on design features of online deliberation platforms with regard to how they address the challenges of scaling up deliberation to crowd size. We found that the most commonly studied specific design features are argumentation tools and facilitation tools. While these tools address certain issues such as the large volume of information and need for structuring arguments, they may neglect the more nuanced requirements of high quality deliberation and possibly reduce inclusion or uptake. They may also inadvertently create new problems relating to inclusiveness or equality. Our analysis of the literature shows that the characteristics (age, gender, education, ethnicity/race, etc.) of test users is rarely reported in case studies we found and the design of deliberation tools mainly takes place in a Western context. Based on our findings, deliberative platforms are more likely to reflect the values and needs of a small, unrepresentative, segment of the world's population. The resulting mismatch with the technologies and social groups that use them risks impeding a wider uptake.

Some designs that feature gamification or allow anonymity or asynchronous participation may address certain issues like building empathy or avoiding some types of discrimination in online platforms. However, in general, much research is still needed on how to facilitate and accommodate different genders or cultures in deliberation, how to deal with the implications of pre-existing social inequalities, how to deal with differences in cognitive abilities and cultural or linguistic differences or how to build motivation and self-efficacy in certain groups. In order to ensure that a fair interpretation of deliberative ideals is reflected in platform design, rather than an expert driven interpretation (Alnemr, 2020), the use of design methodologies such as participative (PD) or value-sensitive design (VSD) is recommended. Design methodologies that involve stakeholder participation are more likely to reflect the needs and values of the intended users. However, the selection of stakeholders must be carefully executed. Universal design (see López and Farzan, 2017) may be particularly interesting, since it involves the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design. Our findings highlight the importance of better integrating approaches from gender studies, psychology, anthropology, psychology into design processes in order to create attractive and inclusive deliberation platforms going forward. It will also be crucial to examine the impact of new design features on the uptake of the software tools as well

as the quality of deliberation. Having said all of this, this should be done bearing in mind that aside from platform design aspects, other external factors, such as political will, cultural or historical reasons may also impact on the uptake of online deliberation.

## Author contributions

Introduction and discussion were primarily developed by RS and AI. Literature review and results section was equally distributed between RS, AI, PM, and MM. CJ offered guidance and supervision. All authors reviewed and proofread the paper. All authors contributed to the article and approved the submitted version.

## Funding

This research was partially funded by the Hybrid Intelligence Center, a 10-year programme funded by the Dutch Ministry of Education, Culture, and Science through the Netherlands Organisation for Scientific Research (Grant number 024.004.022) and by Scientific Research (NWO) Grant-Participatory Value Evaluation: A new assessment model for promoting social acceptance of sustainable energy project (313-99-333). This research was also partially supported by the European Commission funded project-Humane AI: Toward AI Systems that Augment and Empower Humans by Understanding Us, our Society and the World Around Us (grant # 820437).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpos.2022.946589/full#supplementary-material>

## References

- Afsahi, A. (2021). Gender difference in willingness and capacity for deliberation. *Soc. Polit.* 28, 1046–1072. doi: 10.1093/sp/jxaa003
- Alnemr, N. (2020). Emancipation cannot be programmed: blind spots of algorithmic facilitation in online deliberation. *Contemp. Polit.* 26, 531–552. doi: 10.1080/13569775.2020.1791306
- Asenbaum, H. (2016). Facilitating inclusion: austrian wisdom councils as democratic innovation between consensus and diversity. *J. Public Deliberat.* 12, 7. doi: 10.16997/jdd.259
- Bächtiger, A., and Hangartner, D. (2010). When deliberative theory meets empirical political science: theoretical and methodological challenges in political deliberation. *Polit. Stud.* 58, 609–629. doi: 10.1111/j.1467-9248.2010.00835.x
- Beauvais, E. (2018). “Deliberation and equality,” in *The Oxford Handbook of Deliberative Democracy*, eds J. Mansbridge, M. Warren, A. Bächtiger, and J. Dryzek (Oxford: Oxford University Press), 144–155.
- Beck, J., Neupane, B., and Carroll, J. M. (2019). Managing conflict in online debate communities. *First Monday*. 24. doi: 10.5210/fm.v24i7.9585
- Becker, J., Porter, E., and Centola, D. (2019). The wisdom of partisan crowds. *Proc. Nat. Acad. Sci. U.S.A.* 116, 10717–10722. doi: 10.1073/pnas.1817195116
- Bista, S. K., Nepal, S., Paris, C., and Colineau, N. (2014). Gamification for online communities: a case study for delivering government services. *Int. J. Cooperative Information Syst.* 23, 1441002. doi: 10.1142/S0218843014410020
- Black, L. W., Welsch, H. T., Cosley, D., and DeGroot, J. M. (2011). Self-governance through group discussion in wikipedia: measuring deliberation in online groups. *Small Group Res.* 42, 595–634. doi: 10.1177/1046496411406137
- Boschi, G., Young, A. P., Joglekar, S., Cammarota, C., and Sastry, N. (2021). Who has the last word? Understanding how to sample online discussions. *ACM Trans. Web* 15, 1–25. doi: 10.1145/3452936
- Bozdag, E., and Van Den Hoven, J. (2015). Breaking the filter bubble: democracy and design. *Ethics Inf. Technol.* 17, 249–265. doi: 10.1007/s10676-015-9380-y
- Brinker, D. L., Gastil, J., and Richards, R. C. (2015). Inspiring and informing citizens online: a media richness analysis of varied civic education modalities. *J. Comput. Mediated Commun.* 20, 504–519. doi: 10.1111/jcc4.12128
- Cayrol, C., and Lagasque-Schiex, M.-C. (2005). “On the acceptability of arguments in bipolar argumentation frameworks,” in *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty* (Springer), 378–389.
- Chambers, S. (2018). Human life is group life: deliberative democracy for realists. *Crit. Rev.* 30, 36–48. doi: 10.1080/08913811.2018.1466852
- Chesnevar, C., Modgil, S., Rahwan, I., Reed, C., Simari, G., South, M., et al. (2006). Towards an argument interchange format. *Knowl. Eng. Rev.* 21, 293–316. doi: 10.1017/S0269888906001044
- Christensen, H. S. (2021). A conjoint experiment of how design features affect evaluations of participatory platforms. *Gov. Inf. Q.* 38, 101538. doi: 10.1016/j.giq.2020.101538
- Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* 58, 1015. doi: 10.1037/0022-3514.58.6.1015
- Cohen, J. (ed.). (2009). “Reflections on deliberative democracy,” in *Philosophy, Politics, Democracy* (Cambridge, MA: Harvard University Press), 326–347.
- Coleman, S., and Moss, G. (2012). Under construction: the field of online deliberation research. *J. Information Technol. Polit.* 9, 1–15. doi: 10.1080/19331681.2011.635957
- Conklin, J. (2005). *Dialogue Mapping: Building Shared Understanding of Wicked Problems*. Chichester: John Wiley & Sons, Inc.
- Curato, N., Hammond, M., and Min, J. B. (2019). *Power in Deliberative Democracy*. Cham: Palgrave Macmillan.
- Dryzek, J. (2010). *Foundations and Frontiers of Deliberative Governance*. Oxford: Oxford University Press.
- Dryzek, J. S., Bächtiger, A., Chambers, S., Cohen, J., Druckman, J. N., Felicetti, A., et al. (2019). The crisis of democracy and the science of deliberation. *Science* 363, 1144–1146. doi: 10.1126/science.aaw2694
- Durnova, A., Fischer, F., and Zittoun, P. (2016). “Discursive approaches to public policy: politics, argumentation, and deliberation,” in *Contemporary Approaches to Public Policy*, eds B. G. Peters and P. Zittoun (Springer), 35–56.
- El-Assady, M., Gold, V., Hautli-Janisz, A., Jentner, W., Butt, M., Holzinger, K., et al. (2016). “Visargue: a visual text analytics framework for the study of deliberative communication,” in *PolText 2016-The International Conference on the Advances in Computational Analysis of Political Text* (Zagreb), 31–36.
- Epstein, D., and Leshed, G. (2016). The magic sauce: practices of facilitation in online policy deliberation. *J. Public Deliberat.* 12. doi: 10.16997/jdd.244
- Epstein, D., Newhart, M., and Vernon, R. (2014). Not by technology alone: the “analog” aspects of online public engagement in policymaking. *Gov. Inf. Q.* 31, 337–344. doi: 10.1016/j.giq.2014.01.001
- Ercan, S. A., Hendriks, C. M., and Dryzek, J. S. (2019). Public deliberation in an era of communicative plenty. *Policy Polit.* 47, 19–36. doi: 10.1332/030557318X15200933925405
- Faddoul, M., Chaslot, G., and Farid, H. (2020). A longitudinal analysis of youtube’s promotion of conspiracy videos. *arXiv preprint. arXiv:2003.03318*. doi: 10.48550/arXiv.2003.03318
- Fishkin, J. (2020). Cristina lafont’s challenge to deliberative minipublics. *J. Deliberat. Democracy*. 16, 56–62. doi: 10.16997/jdd.394
- Fishkin, J., Garg, N., Gelauff, L., Goel, A., Munagala, K., Sakshuwong, S., et al. (2018). “Deliberative democracy with the online deliberation platform,” in *The 7th AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2019)* (HCOMP).
- Fishkin, J. S. (2009). “Virtual public consultation: prospects for internet deliberative democracy,” in *Online Deliberation: Design, Research, and Practice*, eds T. Davies and S. P. Gangadharan (Stanford: CSLI Publications), 23–35.
- Font, J., Wojcieszak, M., and Navarro, C. J. (2015). Participation, representation and expertise: citizen preferences for political decision-making processes. *Polit. Stud.* 63, 153–172. doi: 10.1111/1467-9248.12191
- Fournier-Tombs, E., and Di Marzo Serugendo, G. (2020). Delibanalysis: understanding the quality of online political discourse with machine learning. *J. Information Sci.* 46, 810–822. doi: 10.1177/0165551519871828
- Friess, D., and Eilders, C. (2015). A systematic review of online deliberation research. *Policy Internet* 7, 319–339. doi: 10.1002/poi3.95
- Fujita, K., Ito, T., and Klein, M. (2017). “Enabling large scale deliberation using ideation and negotiation-support agents,” in *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)* (IEEE), 360–363.
- Gastil, J., and Black, L. (2007). Public deliberation as the organizing principle of political communication research. *J. Public Deliberat.* 4. doi: 10.16997/jdd.59
- Gastil, J., and Broghammer, M. (2021). Linking theories of motivation, game mechanics, and public deliberation to design an online system for participatory budgeting. *Polit. Stud.* 69, 7–25. doi: 10.1177/0032321719890815
- Gold, V., USS, B. P., El-Assady, M., Sperrle, F., Budzynska, K., Hautli-Janisz, A., et al. (2018). “Towards deliberation analytics: stream processing of argument data for deliberative communication,” in *Proceedings of COMMA Workshop on Argumentation and Society*, 1–3.
- Gonçalves, F. M., Prado, A. B., and Baranauskas, M. C. C. (2020). “Opensign: analyzing deliberation and rationale in an exploratory case study,” in *ICEIS*, 511–522. doi: 10.5220/0009385305110522
- Goodin, R. E. (2000). Democratic deliberation within. *Philos. Public Affairs* 29, 81–109. doi: 10.1111/j.1088-4963.2000.00081.x
- Gordon, E., and Manosevitch, E. (2011). Augmented deliberation: merging physical and virtual interaction to engage communities in urban planning. *New Media Soc.* 13, 75–95. doi: 10.1177/1461444810365315
- Gordon, E., Michelson, B., and Haas, J. (2016). “@ stake: a game to facilitate the process of deliberative democracy,” in *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (Princeton, NJ), 269–272.
- Gu, W., Moustafa, A., Ito, T., Zhang, M., and Yang, C. (2018). “A case-based reasoning approach for automated facilitation in online discussion systems,” in *2018 Thirteenth International Conference on Knowledge, Information and Creativity Support Systems (KICSS)* (IEEE), 1–5.
- Gürkan, A., and Iandoli, L. (2009). “Common ground building in an argumentation-based online collaborative environment,” in *Proceedings of the International Conference on Management of Emergent Digital EcoSystems* (New York, NY), 320–324. doi: 10.1145/1643823.1643882
- Gürkan, A., Iandoli, L., Klein, M., and Zollo, G. (2010). Mediating debate through on-line large-scale argumentation: evidence from the field. *Inf. Sci.* 180, 3686–3702. doi: 10.1016/j.ins.2010.06.011



- Gutmann, A., and Thompson, D. F. (2004). *Why Deliberative Democracy?* Princeton, NJ: Princeton University Press.
- Hamari, J., Koivisto, J., and Sarsa, H. (2014). "Does gamification work?—a literature review of empirical studies on gamification," in *2014 47th Hawaii International Conference on System Sciences* (Waikoloa, HI: IEEE), 3025–3034.
- Hartz-Karp, J., and Sullivan, B. (2014). The unfulfilled promise of online deliberation. *J. Public Deliberation* 10, 1–5. doi: 10.16997/jdd.191
- Hassan, L. (2017). Governments should play games: towards a framework for the gamification of civic engagement platforms. *Simul. Gaming* 48, 249–267. doi: 10.1177/1046878116683581
- Hassan, L., and Hamari, J. (2020). Gameful civic engagement: a review of the literature on gamification of eparticipation. *Gov. Inf. Q.* 37, 101461. doi: 10.1016/j.giq.2020.101461
- Hendriks, C. M. (2016). Coupling citizens and elites in deliberative systems: the role of institutional design. *Eur. J. Polit. Res.* 55, 43–60. doi: 10.1111/1475-6765.12123
- Himmelroos, S., Rapeli, L., and Grönlund, K. (2017). Talking with like-minded people—equality and efficacy in enclave deliberation. *Soc. Sci. Jo.* 54, 148–158. doi: 10.1016/j.soscij.2016.10.006
- Iandoli, L., Quinto, I., De Liddo, A., and Buckingham Shum, S. (2016). On online collaboration and construction of shared knowledge: assessing mediation capability in computer supported argument visualization tools. *J. Assoc. Information Sci. Technol.* 67, 1052–1067. doi: 10.1002/asi.23481
- Iandoli, L., Quinto, I., De Liddo, A., and Shum, S. B. (2014). Socially augmented argumentation tools: rationale, design and evaluation of a debate dashboard. *Int. J. Hum. Comput. Stud.* 72, 298–319. doi: 10.1016/j.ijhcs.2013.08.006
- Iandoli, L., Quinto, I., Spada, P., Klein, M., and Calabretta, R. (2018). Supporting argumentation in online political debate: evidence from an experiment of collective deliberation. *New Media Soc.* 20, 1320–1341. doi: 10.1177/1461444817691509
- Isemann, S., and Reuter, W. D. (1997). "Ibis—a convincing concept... but a lousy instrument?" in *Proceedings of the 2nd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (New York, NY), 163–172.
- Ito, T. (2018). "Towards agent-based large-scale decision support system: the effect of facilitator," in *Proceedings of the 51st Hawaii International Conference on System Sciences* (Waikoloa, HI), 351–360.
- Ito, T., Hadfi, R., Haqbeen, J., Suzuki, S., Sakai, A., Kawamura, N., et al. (2020). "Agent-based crowd discussion support system and its societal experiments," in *International Conference on Practical Applications of Agents and Multi-Agent Systems* (Cham: Springer), 430–433.
- Jacquet, V. (2017). Explaining non-participation in deliberative mini-publics. *Eur. J. Polit. Res.* 56, 640–659. doi: 10.1111/1475-6765.12195
- Jhaver, S., Vora, P., and Bruckman, A. (2017). *Designing for Civil Conversations: Lessons Learned From Changemyview*. Atlanta, GA: Technical report, Georgia Institute of Technology.
- Johnson, I. G., MacDonald, A., Briggs, J., Manuel, J., Salt, K., Flynn, E., et al. (2017). "Community conversational: supporting and capturing deliberative talk in local consultation processes," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY), 2320–2333. doi: 10.1145/3025453.3025559
- Jonsson, M. E., and Åström, J. (2014). The challenges for online deliberation research: a literature review. *Int. J. E Polit.* 5, 1–15. doi: 10.4018/ijep.2014010101
- Kennedy, R., Sokhey, A. E., Abernathy, C., Esterling, K. M., Lazer, D. M., Lee, A., et al. (2020). Demographics and (equal?) voice: assessing participation in online deliberative sessions. *Polit. Stud.* 69, 66–88. doi: 10.1177/0032321719890805
- Kim, H., Ko, E.-Y., Han, D., Lee, S.-C., Perrault, S. T., Kim, J., et al. (2019). "Crowdsourcing perspectives on public policy from stakeholders," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY), 1–6. doi: 10.1145/3290607.3312769
- Kirschner, P. A., Buckingham-Shum, S. J., and Carr, C. S. (2012). *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. Berlin: Springer Science and Business Media.
- Klein, M. (2012). Enabling large-scale deliberation using attention-mediation metrics. *Comput. Support. Cooperative Work* 21, 449–473. doi: 10.1007/s10606-012-9156-4
- Kubin, E., Puryear, C., Schein, C., and Gray, K. (2021). Personal experiences bridge moral and political divides better than facts. *PNAS*. 118, e2008389118. doi: 10.1073/pnas.2008389118
- Kunz, W., and Rittel, H. W. (1970). *Issues as Elements of Information Systems*, Vol. 131. Berkeley, CA: Institute of Urban and Regional Development, University of California, p. 14.
- Lafont, C. (2019). Democracy without shortcuts. *Constellations* 26, 355–360. doi: 10.1111/1467-8675.12432
- Lee, S.-C., Song, J., Ko, E.-Y., Park, S., Kim, J., and Kim, J. (2020). "Solutionchat: real-time moderator support for chat-based structured discussion," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY), 1–12. doi: 10.1145/3313831.3376609
- Leighninger, M. (2012). "Mapping deliberative civic engagement," in *Democracy in Motion: Evaluating the Practice and Impact of Deliberative Civic Engagement*, eds T. Nabatchi, J. Gastil, G. M. Weiksner, and M. Leighninger (Oxford: Oxford University Press), 19.
- Levy, K., and Barocas, S. (2017). Designing against discrimination in online markets. *Berkeley Technol. Law J.* 32, 1183–1238. doi: 10.15779/Z38BV79V7K
- Liddo, A. D., and Shum, S. B. (2013). "Improving online deliberation with argument network visualization," in *Digital Cities 8. 6th International Conference on Communities and Technologies* (New York, NY).
- López, C., and Farzan, R. (2017). "Designing for digital inclusion: a post-hoc evaluation of a civic technology," in *International Conference on Social Informatics* (Cham: Springer), 572–588.
- Maia, R. C., and Rezende, T. A. (2016). Respect and disrespect in deliberation across the networked media environment: examining multiple paths of political talk. *J. Comput. Mediated Commun.* 21, 121–139. doi: 10.1111/jcc4.12155
- Manosevitch, E., Steinfeld, N., and Lev-On, A. (2014). Promoting online deliberation quality: cognitive cues matter. *Information Commun. Soc.* 17, 1177–1195. doi: 10.1080/1369118X.2014.899610
- Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Parkinson, J., Fung, A., et al. (2012). *A systematic approach to deliberative democracy*. Deliberative Systems: Deliberative Democracy at the Large Scale. p. 1–41. Available online at: [http://www.cambridge.org/gb/knowledge/isbn/item6796478/?site\\_locale=en\\_GB%5Chttp://www.worldcat.org/title/deliberative-systems-deliberative-democracy-at-the-large-scale/oclc/784577135&referer=brief\\_results](http://www.cambridge.org/gb/knowledge/isbn/item6796478/?site_locale=en_GB%5Chttp://www.worldcat.org/title/deliberative-systems-deliberative-democracy-at-the-large-scale/oclc/784577135&referer=brief_results)
- Mercier, H., and Landemore, H. (2012). Reasoning is for arguing: understanding the successes and failures of deliberation. *Polit. Psychol.* 33, 243–258. doi: 10.1111/j.1467-9221.2012.00873.x
- Min, S. J. (2014). On the westernness of deliberation research. *J. Deliberative Democracy* 10. doi: 10.16997/jdd.207
- Monnoyer-Smith, L., and Wojcik, S. (2012). Technology and the quality of public deliberation: a comparison between on and offline participation. *Int. J. Electronic Governance* 5, 24–49. doi: 10.1504/IJEG.2012.047443
- Muchnik, L., Aral, S., and Taylor, S. J. (2013). Social influence bias: a randomized experiment. *Science* 341, 647–651. doi: 10.1126/science.1240466
- Murray, T., Wing, L., Woolf, B., Wise, A., Wu, S., Clark, L., et al. (2013). "A prototype facilitators dashboard: assessing and visualizing dialogue quality in online deliberations for education and work," in *Submitted to 2013 International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government*. Available online at: <https://dl.acm.org/doi/book/10.5555/2685918>
- Neblo, M. A., Esterling, K. M., Kennedy, R. P., Lazer, D. M., and Sokhey, A. E. (2010). Who wants to deliberate—and why? *Am. Polit. Sci. Rev.* 104, 566–583. doi: 10.1017/S0003055410000298
- Osborne, D. M., Byrne, J. H., Massey, D. L., and Johnston, A. N. (2018). Use of online asynchronous discussion boards to engage students, enhance critical thinking, and foster staff-student/student-student collaboration: a mixed method study. *Nurse Educ. Today* 70, 40–46. doi: 10.1016/j.nedt.2018.08.014
- Peacock, C., Scacco, J. M., and Jomini Stroud, N. (2019). The deliberative influence of comment section structure. *Journalism* 20, 752–771. doi: 10.1177/1464884917711791
- Perrault, S. T., and Zhang, W. (2019). "Effects of moderation and opinion heterogeneity on attitude towards the online deliberation experience," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY), 1–12. doi: 10.1145/3290605.3300247
- Pingree, R. J. (2009). "Decision structure: a new approach to three problems in deliberation," in *Online Deliberations: Design, Research, and Practice*, eds T. Davies, and S. P. Gangadharan (Stanford: CSLI Publications), 309–316.
- Popa, E. O., Blok, V., and Wesslink, R. (2020). Discussion structures as tools for public deliberation. *Public Understand. Sci.* 29, 76–93. doi: 10.1177/0963662519880675
- Ramsey, K. S., and Wilson, M. W. (2009). "Rethinking the 'informed' participant: precautions and recommendations for the design of online deliberation," in *Online*

*Deliberation: Design, Research, and Practice*, eds T. Davies, and S. P. Gangadharan (Stanford: CSLI Publications), 259–267.

Reed, C. (2010). “How dialogues create arguments,” in *ISSA Proceedings*.

Rhee, J. W., and Kim, E. M. (2009). “Deliberation on the net: lessons from a field experiment,” in *Online Deliberation: Design, Research, and Practice*, eds T. Davies, and S. P. Gangadharan (Stanford: CSLI Publications), 223–232.

Rose, J., and Sæbø, Ø. (2010). Designing deliberation systems. *Information Soc.* 26, 228–240. doi: 10.1080/01972241003712298

Ross, L., and Ward, A. (1995). Psychological barriers to dispute resolution. *Adv. Exp. Soc. Psychol.* 27, 255–304. doi: 10.1016/S0065-2601(08)60407-4

Rossini, P., Stromer-Galley, J., Baptista, E. A., and Veiga de Oliveira, V. (2021). Dysfunctional information sharing on whatsapp and facebook: the role of political talk, cross-cutting exposure and social corrections. *New Media Soc.* 23, 2430–2451. doi: 10.1177/1461444820928059

Ruckenstein, M., and Turunen, L. L. M. (2020). Re-humanizing the platform: content moderators and the logic of care. *New Media Soc.* 22, 1026–1042. doi: 10.1177/1461444819875990

Sandover, R., Moseley, A., and Devine-Wright, P. (2020). *Rapid Review: The Feasibility of an Online Citizens' Assembly to Support Devon's Transition to Net Zero*. Available online at: <https://www.devonclimateemergency.org.uk/wp-content/uploads/2020/12/Rapid-Review-Online-Deliberation.pdf>

Sarmiento, R., and Mendonça, R. F. (2016). Disrespect in online deliberation: Inducing factors and democratic potentials. *Revista de Ciencia Política (Santiago)*. 36, 705–729. doi: 10.4067/s0718-090x2016000300005

Schleifer, D., and Diep, A. (2019). *Strengthening Democracy: What Do Americans Think*. New York, NY: The 2019 Yankelovich Democracy Monitor Report (Public Agenda, August 2019).

Semaan, B., Faucett, H., Robertson, S. P., Maruyama, M., and Douglas, S. (2015). “Designing political deliberation environments to support interactions in the public sphere,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY), 3167–3176.

Shi, F., Teplitskiy, M., Duede, E., and Evans, J. A. (2019). The wisdom of polarized crowds. *Nat. Human Behav.* 3, 329–336. doi: 10.1038/s41562-019-0541-6

Shortall, R. (2020). “Deliberative appraisal methods,” in *Advances in Transport Policy and Planning, Vol. 7, 1st edn*, ed N. Mouter (Elsevier).

Siu, A. (2017). Deliberation and the challenge of inequality. *Daedalus* 146, 119–128. doi: 10.1162/DAED\_a\_00451

Steenberger, M., Bächtiger, A., Spöndli, M. and Steiner, J. (2003). Measuring political deliberation: a discourse quality index. *Comparat. Euro. Polit.* 1, 21–48. Available online at: <https://content.csbs.utah.edu/~burbank/steenbergen2003.pdf>

Strandberg, K., and Grönlund, K. (2012). Online deliberation and its outcome-evidence from the virtual polity experiment. *J. Inf. Technol. Politics* 9, 167–184. doi: 10.1080/19331681.2011.637709

Stromer-Galley, J., Webb, N., and Muhlberger, P. (2012). Deliberative e-rulemaking project: challenges to enacting real world deliberation. *J. Inf. Technol. Politics* 9, 82–96. doi: 10.1080/19331681.2012.635971

Suiter, J., Muradova, L., Gastil, J., and Farrell, D. M. (2020). Scaling up deliberation: testing the potential of mini-publics to enhance the deliberative capacity of citizens. *Swiss Polit. Sci. Rev.* 26, 253–272. doi: 10.1111/spr.12405

Toots, M. (2019). Why e-participation systems fail: the case of Estonia's osale.ee. *Gov. Inf. Q.* 36, 546–559. doi: 10.1016/j.giq.2019.02.002

Towne, W. B., and Herbsleb, J. D. (2012). Design considerations for online deliberation systems. *J. Information Technol. Politics* 9, 97–115. doi: 10.1080/19331681.2011.637711

Ullmann, S., and Tomalin, M. (2020). Quarantining online hate speech: technical and ethical perspectives. *Ethics Inf. Technol.* 22, 69–80. doi: 10.1007/s10676-019-09516-z

Van Eemeren, F. H., Grootendorst, R., Johnson, R. H., Plantin, C., and Willard, C. A. (2013). *Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments*. New York, NY: Routledge.

Velikanov, C. (2017). “Can deliberative governance become inclusive?” in *Proceedings of the 18th Annual International Conference on Digital Government Research* (New York, NY), 531–540. doi: 10.1145/3085228.3085237

Verdiesen, I., Dignum, V., and Hoven, J. V. D. (2018). Measuring moral acceptability in e-deliberation: a practical application of ethics by participation. *ACM Trans. Internet Technol.* 18, 1–20. doi: 10.1145/3183324

Wironen, M. B., Bartlett, R. V., and Erickson, J. D. (2019). Deliberation and the promise of a deeply democratic sustainability transition. *Sustainability* 11, 1023. doi: 10.3390/su11041023

Wright, S., and Street, J. (2007). Democracy, deliberation and design: the case of online discussion forums. *New Media Soc.* 9, 849–869. doi: 10.1177/1461444807081230

Wulczyn, E., Thain, N., and Dixon, L. (2017). “Ex machina: personal attacks seen at scale,” in *Proceedings of the 26th International Conference on World Wide Web* (Geneva), 1391–1399.

Wyss, D., and Beste, S. (2017). Artificial facilitation: promoting collective reasoning within asynchronous discussions. *J. Inf. Technol. Politics* 14, 214–231. doi: 10.1080/19331681.2017.1338175

Yang, C., Gu, W., Ito, T., and Yang, X. (2021). Machine learning-based consensus decision-making support for crowd-scale deliberation. *Appl. Intell.* 51, 4762–4773. doi: 10.1007/s10489-020-02118-z

Yang, C., Ito, T., and Gu, W. (2019). “Toward machine learning-based facilitation for online discussion in crowd-scale deliberation,” in *Proceedings of the Annual Conference of JSAI 33rd Annual Conference, 2019* (Tokyo: The Japanese Society for Artificial Intelligence), 3H3E304.

Zhang, A. X. (2017). “Systems for improving online discussion,” in *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology, UIST '17* (New York, NY: Association for Computing Machinery), 111–114.

Zhang, W. (2010). “Simulating the ideal e-deliberation: The roles of inclusion, equalization and rationalization,” in *Online Deliberation*, eds F. De Cindio, A. Macintosh, and C. Peraboni (Leeds: OD2010), 117.