Original Article

Strengthening democracy through bottom-up deliberation: An assessment of the internal legitimacy of the G1000 project

Didier Caluwaerts[a,*] and Min Reuchamps[b]

[a]Harvard University, Ash Center for Democratic Governance and Innovation, 79 JFK Street, Mailbox 74, Cambridge, MA 02138, USA.

[b]Université catholique de Louvain, Institut de sciences politiques Louvain-Europe (ISPOLE), Place Montesquieu 1 bte L2.08.07, Louvain-la-Neuve, 1348, Belgium.

*Corresponding author.

Abstract Recent scholarship claims that citizen deliberation can contribute to the quality of democracy and to the legitimacy of political decision making. By including everyone who is affected by a decision in the process leading to that decision, deliberation is capable of generating political decisions that receive broad public support, even when there is strong disagreement on the values a polity should promote. However, if deliberative democracy wants to contribute to the legitimacy of the political system, it has to be legitimate in itself. In other words, deliberative processes have to reflect the principles of legitimacy in their own functioning. It is therefore crucial to assess the internal legitimacy of deliberative mini-publics before making claims about their contribution to the legitimacy of the political system as a whole. In this article, we set out to refine the theory on deliberative legitimacy and to determine the legitimacy of one particularly interesting deliberative event, namely the Belgian G1000. We will argue that it is very difficult for deliberative processes to be high on all dimensions of legitimacy and that there is a trade-off between input and output legitimacy. Moreover, we find that design characteristics to a large extent determine the legitimacy of deliberative processes.

Introduction

From the beginning of the 1990s onwards, all Western European countries witnessed the contours of a widespread crisis of democracy. The alleged decline of political trust and the rise of electoral volatility pointed out that the gap between politicians and citizens had never been wider. This political climate characterized by a

deep-rooted crisis of democratic legitimacy offered an excellent breeding ground for critical reflection on the role and shape of democracy in modern societies. It gave rise to a fruitful quest for new and innovative ways of governing a democracy.

It is in this turbulent period that the ideal of a deliberative democracy was coined. A community of international scholars, inspired by the work of Jürgen Habermas, argued that a vibrant democracy is more than the aggregate of its individual citizens. The quality of democratic decisions, according to them, did not depend on the correct aggregation of individual preferences, but rather on the quality of the public debate that preceded the voting stage (Dryzek, 2000). The quality of democratic decisions was thus determined by extensive argumentation about political choices before voting on them.

This deliberative model started out in life as a theory of legitimacy (Cohen, 1997; Dryzek, 2001; Parkinson, 2006). By including everyone affected in the process leading to a decision, deliberation could prove capable of generating political outcomes that receive broad public support, even when there is strong disagreement on the aims and values a polity should promote.

However, these beneficial effects do not come about easily. If deliberative democracy wants to contribute to increasing the legitimacy of the political system as a whole, it has to be legitimate in itself. In other words, deliberative processes have to reflect the principles of legitimacy in their own functioning before their outcomes can generate legitimate political decisions. It is therefore crucial to assess the internal legitimacy of deliberative mini-publics before making claims about their contribution to the legitimacy of the political system as a whole. Our research question is therefore: to what extent can deliberative mini-publics live up to the criteria of democratic legitimacy?

In this article, we set out to assess the internal legitimacy of one specific deliberative event, namely the G1000 Citizens' Summit in Belgium. The G1000 project takes a particular place in the world of deliberative practice because it was completely grassroots in its process and its organization (Caluwaerts and Reuchamps, 2012). Most deliberative events are set up by either public administrations or scientific institutions, whereas the G1000 was a citizens' initiative from its very inception. All of the organizers of the event were volunteers, and all of the funds were gathered using crowd funding (G1000, 2012). This grassroots structure gives it a remarkable place in the deliberative firmament, and it makes it a very interesting case for students of legitimacy.

In what follows, we will first develop how we can measure a deliberative event's internal legitimacy. Afterwards we analyze how the G1000 was organized and what its design features are. Finally, we assess the internal legitimacy of the project.

Measuring the Internal Legitimacy of Deliberative Processes

In one of the most cited articles on deliberative legitimacy, Cohen (1997) contends that 'outcomes are democratically legitimate if and only if they could be the object of

free and reasoned argument among equals' (p. 74). Like most deliberative scholars (Manin, 1987; Dryzek, 2001; Parkinson, 2006), Cohen thus offers a proceduralist view that links inputs from citizens to political outputs through a certain deliberative procedure.

Building on previous studies on democratic legitimacy, Bekkers and Edwards (2007) and Edwards (2007) discern three dimensions of deliberative legitimacy, which are closely related to Cohen's definition, as well as more generic theories of democratic legitimacy. Democratic decision-making procedures, in their opinion, have to be legitimate in the input, throughput and output phases: they have to make sure that the opinions and needs of ordinary citizens are translated through deliberative procedures into good political outcomes. This threefold model dates back to Easton's (1965) seminal work on political systems and is often used in studies on democracy as such (see, for example, Papadopoulos and Warin, 2007; Schmidt, 2013), that is, without the deliberative attribute. Although Easton's model was geared more strongly toward traditional representative decision-making processes, Bekkers and Edwards (2007) adapted it to be applicable to new democratic repertoires. This means that it is inspired by general theories of legitimacy, but tailored to fully capture the specificity of innovative democratic practices such as deliberative democracy (Hendriks, 2010). We discuss each of these three dimensions of deliberative legitimacy hereafter.

Input legitimacy

The input legitimacy of deliberative events deals with citizens' opportunities to influence the process and the outcomes of the deliberation. Input legitimacy is thus a measure for the openness of the deliberative events toward demands and needs from its participants. It is high if citizens have the chance to fully make their opinions known. Inclusiveness is thus the central principle, whereas selectivity has to be avoided in order to get a genuine reflection of the authentic preferences of the population at large. According to Edwards (2007), this type of legitimacy has to meet two central criteria in a deliberative setting, namely, the quality of representation and the openness of the agenda.

Quality of representation

One of the main problems for deliberative democracy is the scale problem. The kind of argumentative interaction between citizens that deliberation entails is very difficult to achieve in mass democracies, and 'meaningful participation in collective decision by anything more than a tiny minority is inconceivable in contemporary nation-states' (Dryzek, 2001, p. 652). As such, in order to make citizen deliberation viable, it is usually scaled down to a mini-public, taking into account that the participants are in some way representative of the larger population.

For this selection process, forms of descriptive representation through random sampling are considered to be ideal (Landemore, 2012), because they offer each citizen an equal chance of being part of the mini-public, and they increase cognitive diversity. Rather than being selectionistic and choosing only the most able citizens, deliberative organizers should include an epistemically diverse set of participants (Page, 2007). This allows for a thorough process of argumentation in which all public positions are represented, and which eventually leads to more legitimate decisions (Thompson, 2008). After all, only when all ideas are heard, can the best idea be identified.

To guarantee such an experiential diversity, random selection is often considered to be the most appropriate technique (Fishkin *et al*, 2000). In contrast, sampling techniques that foster self-selection, such as snowball sampling or very open calls for participation, only foster inbreeding among likeminded and therefore undermine the input legitimacy (Ryfe, 2005).

Openness of the agenda

Besides a high quality of representation, legitimate deliberation also requires an open agenda. If the agenda is closed, the diversity in the group will not be able to manifest itself, because citizens' opinions will be restricted to a very narrow set of items. Selectivity during the agenda setting will thus undermine the input legitimacy because issues are banned from the discussion (Edwards, 2007). As policy problems are often holistic and transversal, a closed agenda setting and very narrowly defined topics hinder the inclusivity of the event on the input side. Allowing the participants to explore new and adjacent problems on the other hand could increase the input legitimacy.

Throughput legitimacy

Whereas the input phase deals with the inclusion of participants and their ideas in the deliberation, the throughput phase focuses on the deliberative process itself. The procedures used to shape the process of deliberation determine to a large extent the quality of that deliberation and its outcome. It is therefore crucial to assess the throughput legitimacy based on the following three criteria.

Quality of participation

It is not sufficient that different voices are represented in a mini-public; these voices also have to be heard. The deliberation has therefore to be modeled in such a way as to bring out everyone's experience and perspectives (taking into account their different ability to express their views), and to foster openness toward the arguments of others. In this respect, the quality of participation captures whether the discussion reflects the characteristics of the ideal speech situation (Bekkers and Edwards, 2007).

Crucial determinants of the quality of participation are therefore the style of moderation (active involvement of the moderator can increase equal participation), the setting (hot deliberative settings are more confrontational than cold ones and therefore scare more people off), the size of the groups (large and diverse groups are more difficult to manage and have exclusionary tendencies) and other inhibiting factors such as the use of multiple languages (Caluwaerts and Deschouwer, 2013; Caluwaerts and Reuchamps, 2014).

Quality of decision making

It is important that decisions come about through argumentation, and not through coercion. They must reflect the reasoned opinion and openness to persuasion of all those involved and not the power relations in the group. However, the theoretical aim of uncoercive consensus is at odds with the real world of politics, where it is inevitable that conflicts will continue to linger. 'Even under ideal conditions', Cohen (1997, p. 75) contends, 'there is no promise that consensual reasons will be forthcoming. If they are not, then deliberation concludes with voting, subject to some form of majority rule'.

This makes clear that there is a place for voting in deliberative democracy (Saward, 2000, pp. 67–68). Most deliberative designs therefore rely on some aggregative mechanisms to come to some kind of final decision, so that the power of the majority does play a role. The quality of decision making therefore depends on the number of times aggregation crosses the argumentative process, and on the binding power of these aggregative decisions. For instance, if there is a deliberative process, but it is interrupted every hour with a straw poll to see whether and how opinions have shifted, the power of the majority weighs more on the process than with simply a final vote at the end of the event.

Contextual independence

Deliberation never takes place in a societal vacuum. It is embedded in a certain political context, and this context always influences the process of deliberation. This wider public involvement strengthens deliberation, precisely because the issues are so politically salient, and because the deliberators are always part of that population. A vibrant deliberative democracy should therefore be able to handle these outside influences.

However, at times, participants can experience large outside pressures that undermine deliberation. When they are forced to choose a particular course of action, they suffer from coercion and the process of reasoned argument, as a basis for legitimate decisions, is completely undermined. Deliberators should thus be substantively independent from these outside pressures exerted by political parties, public opinion makers, pressure groups or the media.

Output legitimacy

Finally, deliberative events also have to live up to the legitimacy requirements on the output side. This means that three specific criteria have to be met.

Public endorsement

Because of the scale problem, only a small part of the citizenry was implicated in the deliberative decision-making process, and the outcomes did not receive assent from the wider public. Hence, any deliberative event should have feedback loops to the public as a whole. As Dryzek (2001, p. 654) puts it: 'decisions still have to be justified to those who did not participate'. This means that the results from a deliberative endeavor should be put to the test of publicity once more, and receive public assent. This could be done by putting the results to the popular vote in a referendum, but to some extent feedback might also be organized through more informal means such as televising the deliberations.

Weight of the results

In order for the outputs to be legitimate, they have to be in some way linked to formal political decision making (Edwards, 2007). Otherwise deliberation is merely a form of democratic experimentation without any practical use. The output legitimacy is therefore good when the effective impact of the deliberative outputs on real-world politics is high. This also implies that the output legitimacy increases as influence on formal political processes grows: outcomes with a mere agenda-setting function have less weight than when government has expressed its commitment to implementing the final decisions.

Responsiveness and accountability

Decisions taken through deliberation should also be responsive. This means that the solutions proposed and decisions taken should offer an answer to the problems that were initially identified, and fit with the cultural and value schemes of the wider society in which they have to be applied.

Moreover, there should also be regular feedback to the participants (Bekkers and Edwards, 2007). Those who put their heads together to come up with solutions should be kept in the loop on what happens with these solutions and what changes were made. This means that there should be a transparent chain of responsibility. It should be clear to the participants in deliberative events who can be held accountable for the results that come out of the deliberations.

The G1000 Project

In order to test the potential for generating legitimacy through deliberation, we will study the G1000 (2012) Citizens' Summit. The G1000 is a large-scale deliberative

project that was launched in Belgium in June 2011, and it takes a particular place in the world of deliberative experiments. Unlike most deliberative mini-publics which are organized for research purposes or funded by government institutions, the G1000 was a completely grassroots organization. It relied on – 800 – volunteers for everything from marketing to funding, and the event was not financed by universities or government institutions. Through crowd funding, that is, small donations from private citizens or companies, the organizers succeeded in gathering the budget (€300 000) in less than 6 months.

The reason why the organizers did not apply for public money is because they opted for a very open agenda setting. This was important because the main argument for organizing the Citizens' Summit was the generally held belief that there is an ever-growing gap between politics and citizens, and that the public and political agendas no longer coincide (Caluwaerts, 2011; Reuchamps and Caluwaerts, 2013a). Accepting funds from public institutions would thus prioritize the political agenda once more. Instead, the G1000 aimed at creating a setting, which is conducive to open deliberation on contentious political issues that the citizens themselves found important.

In order to do so, the G1000 project consisted of three distinct – but interrelated – phases, namely, a public consultation, a citizen deliberation and policy preparation. The first phase consisted of a process of agenda setting through public consultation. Contrary to what is a commonplace practice in deliberative ventures, the organizers were convinced of the importance of starting with a very open agenda. The organizers therefore launched an online idea-box that every citizen, regardless of his opinion or background, could post the fundamental questions that should be treated by a citizen forum. This online consultation resulted in a total of over 2000 ideas, which were subsequently clustered into a top 25 of themes based on the number of times they appeared and on their rating. This list of 25 was once again put online in October 2011, and citizens were invited to vote for the three items they wanted to see discussed at the citizen summit. Eventually, these three issues turned out to be: social security, welfare in times of economic crisis and immigration.

After the phase of public consultation, the second part of the project consisted in a process of citizen deliberation, the G1000 itself. On 11 November 2011, sitting at 81 tables, 704 participants were invited to reflect, discuss and argue their positions on the three issues that were put on the agenda. Each table was moderated by a trained facilitator. At the bilingual tables simultaneous translation was foreseen, and the organizers invited a team of international experts to guard over the quality of the deliberative process.

The Citizens' Summit in Brussels was also flanked by two side projects: G'Home and G'Offs (G1000, 2012). The G'Home was a software application allowing citizens to log onto an online discussion, whereas the G'Offs gathered citizens to discuss the same issues as in Brussels but at local tables all over Belgium. There was thus a much larger group than those gathered in Brussels discussing the three main issues.

The third phase of the G1000 project, which is called the G32, tried to further develop the ideas from the G1000 into thought-through policy proposals. The 32 participants gathered during three weekends to hear experts, and to write down policy. This design is much more intensive and requires much more skill from the moderators in order to enable participants to propose specific policies and actions. It is also more open than a citizen summit since the participants have a much greater say in the process itself.

Assessing the Legitimacy of the G1000 Project

The G1000 started out in life as a way of increasing the legitimacy of the Belgian democratic system by giving it a strong bottom-up impulse. By gathering ordinary citizens to speak and reflect on three crucial political issues, it attempted to close the gap between politicians and public, and to come to effective and broadly carried policy proposals (Reuchamps and Caluwaerts, 2013b). However, in order to make claims about its impact, we should assess the legitimacy of the G1000 itself, and see whether it lived up to its claims to direct citizen involvement in its own functioning. We will do this in the next sections.

The analysis relies primarily on the final report of the G1000 (2012), which also contains the report of the international observers. Moreover, in addition to a pre- and post-test with the participants, we also conducted interviews with academics and a selection of the participants, and we organized a survey among politicians as well as among a representative sample of 1000 Belgians to measure their respective perceptions of the G1000.

Input legitimacy

With regard to the input legitimacy, we will look at the quality of the representation and the openness of the agenda in order to assess how well the diverse societal inputs reached the deliberative forum, and whether some views were excluded by the G1000 design.

Quality of representation
The key issue for the input legitimacy of any deliberative event concerns the selection of the participants. Normatively, the most appealing technique for recruiting participants is random selection (Bohman, 2007, pp. 351–352; Reuchamps, 2011; Caluwaerts and Ugarriza, 2012). Randomization ensures that the multitude of public opinions is present in a group and it thus 'produces discussion among people who think and vote differently and would not normally be exposed to one another' (Fishkin *et al*, 2000, p. 660).

This is also the reason why the G1000 opted for random selection. The recruitment procedure aimed at maximizing the diversity of opinions among the participants, in order to avoid 'informational inbreeding among likeminded citizens' (Huckfeldt, 2001, p. 426). Citizens can, after all, only find themselves in a situation of genuine deliberation when they are faced with competing claims (Thompson, 2008). When everyone at the table shares the same opinion, deliberation cannot live up to its claims of input legitimacy.

The participants at the G1000 were recruited through Random Digit Dialing. This technique generates random phone numbers for fixed and mobile lines and has a penetration rate of 99 per cent in Belgium. Every inhabitant thus had an equal chance of being selected for participation. In order to guard over the quality of the sample, the random selection was checked for certain predefined population quotas, among which gender, age and province as to guarantee a proportional representation of both linguistic groups.

In the end, these quotas seem to be well respected in the group of final participants. Fifty-two per cent of the participants were female, 48 per cent were male, which is a perfect reflection of the gender composition of the population, and which was rather unexpected since women are found to be more likely to drop out of such deliberative events (Ryfe, 2005). Moreover, 61 per cent of the participants were Dutch-speaking versus 39 per cent of the French speakers, which is also an accurate reflection of the population. The G1000 thus seems to score rather well when it comes to the quality of the representation.

However, these figures do not tell the entire story because the Citizens' Summit witnessed processes of self-selection and last minute dropout. With regard to self-selection, we should note that the response rate was very low. Only around 3 per cent of those invited eventually said yes, so that the commitment required for participating in a deliberative event proved to be very high. Moreover, the self-selection was much higher within minority groups. The organizers therefore decided to reserve 10 per cent of the seats for participants who were contacted through grassroots organizations engaged in helping people with a lower socioeconomic status. All in all, 90 per cent of the participants were thus randomly selected, whereas 10 per cent came from a targeted recruitment. Such design aimed at maximizing the inclusion of different perspectives, but at the same time it made it much harder to assess the overall representativeness of the sample.

In addition, the event experienced a dropout rate of about 30 per cent among the people who had previously confirmed their participation. The final number of participants therefore amounted up to 704. These dropout effects have to be put into perspective. Their rate is comparable to that of other mini-publics even though the participants of the G1000 did not receive any financial compensation for their participation, which is commonplace in deliberative practice (Ryfe, 2005).

The quality of representation thus requires a qualified assessment. On the one hand, the organizers did rely on random selection in order to get a maximum

diversity of opinions and perspectives around the tables, and the final participant sample was sociodemographically representative of the entire population. On the other hand, the G1000 lost a lot of citizens along the way. There were self-selection biases during the recruitment, and there was a 30 per cent dropout before the event. This limits the quality of representation and the legitimacy of the event on the input side, but these losses of input legitimacy are not unique to the G1000. Every deliberative event experiences them so that the overall assessment of the quality of representation is positive.

Openness of the agenda
In addition to the input-oriented recruitment process, the organizers of the G1000 opted for an open agenda-setting process. Because of the nationalist tensions between the north and the south of the country, it was unwise for the organizers to set the agenda themselves. They therefore organized an online agenda-setting application. Every person living in Belgium could log on to the Website and launch the idea he or she wanted to see discussed at the G1000. This application was sided by a large-scale media campaign to ensure that as many people as possible found their way to the idea box.

In total, over 2000 ideas were launched, and over 6000 people visited the Website to rate the ideas posted by others. Of course, most of these ideas appeared multiple times, which is why they were subsequently clustered into a top 25. This top 25 was once again put on the Website and, through a media campaign, the public was asked to rate their top 3 issues. These issues would form the agenda of the Citizens' Summit.

The agenda-setting process was thus very open, and the aim was to let the public (not the political elite) decide what was going to be discussed at the G1000. The Citizens' Summit relied entirely on the direct inputs from the citizens on which issues had to be discussed, which is why the project scores high on input legitimacy. However, despite this relatively straightforward procedure of clustering the items, we should note that the international observers had one critical remark: 'the process of framing, summarizing and clustering the 2000 proposals to 25 needs to be made more transparent and the methodology should be explained' (G1000, 2012, p. 105).

The high level of input legitimacy also resulted from another specific design characteristic of the event, namely its funding. Unlike many of its international counterparts, the G1000 was entirely financed using crowd funding. It did not accept any money from government for fear that this would narrow the agenda. This is not to say that government-initiated events do not display any openness for the participants to address issues of their concern in a more holistic manner, but they rarely allow outsiders to determine the fundamental themes a mini-public should address. This grassroots characteristic was praised by the international observers who stated that 'the G1000 team secured their independence of financial supporters by ceiling the financial support of an individual actor to 7 per cent of the total budget'

(G1000, 2012, p. 105). The agenda was thus open and the discussions were not shaped to fit the agenda of pressure groups or government.

Throughput legitimacy

The throughput phase deals with the question how the inputs from the participants are treated, and which dynamics guide the decision-making process. We will therefore look at the quality of the decision making, the quality of the participation, and whether the political and societal environment influenced the process of deliberation from the outside.

Quality of decision making

At multiple times during the event, the organizers asked concrete inputs from each of the tables (Caluwaerts and Reuchamps, 2012). Sometimes this was a mere Post-it with feelings or keywords, but most of the time the organizers asked for specific proposals or ideas, and more importantly also the arguments behind these proposals. Every argument that was aired during the discussions was noted down, and equal consideration was given to each participant's perspective. As a consequence a high level of throughput legitimacy characterized the process.

All the ideas that were discussed at the tables were subsequently sent to the central desk. This central desk consisted of six experienced practitioners of such deliberative design who collected the data from each of the tables, and who clustered the proposals from all of the tables. These clusters were resubmitted to the individual vote of the participants at the end of each discussion round. The decision making at the G1000 was thus characterized by an alternation between deliberation and aggregation. According to Bekkers and Edwards (2007), such a combination significantly improves the throughput legitimacy of the process because it combines the substantive depth and elaboration of citizen deliberation with the clarity of a final vote.

What is problematic from a legitimacy point of view, however, is that part of the aggregation was done behind the scenes. After all, the experts at the central desk were asked to do a first clustering of the inputs from the tables in order to see which ideas were introduced and to facilitate the voting round. As such, some genuinely innovative and original ideas, which only appeared once or twice, did not make it to the final vote, and the post-test questionnaire indicated that some of the participants felt like their opinions were not taken seriously, because their ideas were not put to the vote.

The clustering by the central desk, although this is a common practice in any deliberative event (Elliott *et al*, 2005), thus negatively affected the throughput legitimacy. There was a loss of information during the process, even though the final decision-making power was with the participants. In order to increase the throughput legitimacy of the G1000, the clustering by the central desk should therefore have been more transparent.

The throughput legitimacy was moreover compromised by the very busy agenda of the G1000. In only one day, three complex societal problems (social security, immigration and the financial crisis) had to be discussed. This left too little time for thorough argumentation and information exchange. The groups could therefore not dig very deeply into the issues under discussion, and the process of social learning, which is crucial to deliberation, was somewhat hindered. These time restrictions thus further constrained the quality of the decision making.

Quality of participation

As we have seen before, the quality of representation in the G1000 project was high. However, this does not automatically mean that the quality of the participation was of an equally high standard. Deliberative events can have excellent samples, but poor participation simply because some participants are afraid to speak up. Managing the group dynamics is thus imperative for bringing out the 'wisdom of crowds' (Surowiecki, 2004).

In order to give everyone the liberty to utter his or her opinion, the organizers relied on a clear script with strict time slots, and uniform instructions for the moderators, who had moreover received an intensive training beforehand. In order to lower the threshold for participation, the event started with an introductory round so that the participants had an opportunity to get to know each other. Group deliberation on polarized political issues is after all a very unusual and sometimes frightening setting, with which some feel more at ease than others. It was therefore crucial for the quality of participation that the participants got acquainted with each other and that they lost the fear of uttering their opinions in public as research on focus groups claims (Krueger, 1998).

After the introductory round, the participants discussed the three issues at their tables. The central aims of these discussions were to come to clear problem definitions, to suggest solutions and to take stock of the different ideas and perspectives the participants had on the three themes on the agenda. At the beginning of each round, two independent experts – one from each side of the linguistic border – introduced the theme. These experts offered their takes on the problems and proposed possible solutions, so that the access to all relevant information – which is necessary for a full participation – was available.

Nevertheless, as the international observers rightly pointed out, these experts might not have represented the full spectrum of perspectives on the issues at stake (G1000, 2012, pp. 103–104). They mostly offered rather left-wing views on the issues at stake, which could potentially bias the quality of the participation, because experts are usually considered to be authoritative figures, and their one-sided opinions might have muted diverging opinions among the participants. This lowers the quality of the participation somewhat.

However, this was countered by a very structured interaction script. The script was construed in such a way as to maximize the inclusion of all and the process of

information sharing, and to minimize the social thresholds for fully participating in the discussions. The design of the G1000 made sure that the moderators supported the airing of new and even controversial opinions among the participants. Moreover, they also used different interaction styles in order to lower the threshold for full participation and in order to support the natural group dynamics. Sometimes the participants had to discuss in pairs of two, at other times, they discussed the issues with everyone at the table. Sometimes the tasks were very simple like expressing their feelings toward an issue such as immigration, but at other times, the expected outcomes were more complex and the cognitive investment was much larger (Caluwaerts and Reuchamps, 2012). This variation in interaction styles highly increased the quality of the participation because it took away group dynamic thresholds and allowed every participant to utter his or her opinion in a relatively safe environment.

Even though the participants' evaluation of the G1000 made clear that the organizers expected a lot of input and effort from the participants in very – maybe too – little time, it was this built-in alternation between complex and easy tasks, which avoided silencing certain participants, and which made the discussions very effective in bringing out the cognitive diversity of the group.

Contextual independence

The G1000 was embedded in a very specific social and political context. Indeed, the project was initiated when the negotiations between the party leaders of the north and the south of the country were completely deadlocked (Deschouwer and Reuchamps, 2013; Caluwaerts and Reuchamps, 2014). Bringing citizens together in such a controversial political climate has consequences for the contextual independence of the event because the media, politicians and public opinion makers at the time tended to frame the event in such a way as to influence its impact.

Because of its timing and the idea of bringing Dutch-speakers and French-speakers (and a few German-speakers) together, the G1000 suffered from a strong nationalist headwind, especially in the north of the country. Many public opinion makers framed the project as a pro-Belgian event, thereby delegitimizing it in Flanders (van den Broeck, 2011a). Flemish political parties were openly polarized over the project with some being highly supportive of citizen involvement in politics, whereas others were highly critical of the alleged pro-Belgian image of the organizers. The same polarization also took place among public opinion makers who either heralded the G1000 as a democratic innovation, or stated that the project was an open attack on representative democracy with the explicit aim of delegitimizing traditional electoral politics and destabilizing the political system (Rondas, 2011).

The context thus had a big impact, and the political elites and the media were very active in paving the way for their preferred course of action. However, because the agenda setting was open, the topics to be discussed were not determined by any

political agenda. In fact, the three issues that received most votes were not related to the problems between Flemings and Walloons, which were most politically salient at that time. Consequently, while the project and the participants were strongly influenced by the political context, the deliberative design proved fairly robust in neutralizing the political turmoil.

Output legitimacy

After decisions are made, they have to be fed back to the population as a whole and subsequently implemented before they can generate any lasting outcome. In this section, we look at what became of the results of the G1000, and how the political elites were held accountable.

Public endorsement

Good deliberative democracy entails that the decisions made the mini-public should feed back to the maxi-public from which it was drawn. In other words, the mini-public should keep contact with the population, and decisions have to be sanctioned in some way by this larger population. However, it is unlikely that its decisions will be put to any formalized vote because of Belgium's bad experiences with referenda on polarizing issues. This does not mean the organization did not try to create these kinds of feedback loops between the mini-public and the wider citizenry. For instance, it was not only the randomly chosen 704 people in Brussels who deliberated the issues, but also a much wider group of ordinary citizens. The G1000 was after all flanked by the so-called G'Home (an online deliberative tool) and G'Off's (a few dozen local discussion tables). In order to ensure the connection between mini-public and public, everything was put into place in order to ensure that the wider public could participate in parallel discussions all over the country. For instance, the event in Brussels with the presentations of the experts was live streamed on the Website, and the scripts were available for local organizers and moderators.

Despite all good intentions, however, the G1000 did not succeed in finding strong legitimation of its decisions from the wider public. Contrary to other cases with strong endorsements, such as the British Columbia Citizens' Assembly that ended in a referendum (Fournier *et al*, 2011), the G1000 only organized a media campaign. This is a mere one-way stream of information, in which the citizens became informed, but in which they could not actually endorse or disapprove the conclusions. This lack of wider public endorsement strongly weakens the impact of the decisions. As such, the G1000 scores poorly on this first dimension of output legitimacy by organizing insufficient feedback to the wider public.

Weight of the results

Decisions generated through deliberation should also be implemented if they want to be legitimate on the output side. The results of the citizen deliberation have to be given political weight in the policy-making process. It is on this dimension of legitimacy that the G1000 has fallen short. The G1000 grew as a truly bottom-up initiative, in the sense that it was initiated and funded entirely by citizens. This means that there was no commitment whatsoever from the political elites to take up the results from the G1000.

Moreover, as is common in other initiatives (Smith, 2009), the G1000 was looked at with great skepticism by political parties. It took place after more than 500 days of government negotiations, in a political climate that was very hostile to anything that even remotely referred to the linguistic tensions. It thus received a very strong political headwind because it was seen as anti-political and anti-party (see, for example, Rondas, 2011; van den Broeck, 2011a, b; Szoc and Lechat, 2012). This all meant that, apart from the presidents of the different Belgian parliaments saying that it was important to listen to citizens, very little specific action was taken by politicians to apply the ideas of the G1000 in the policy process.

However, the lack of weight was not only because of the bottom-up organization of the G1000 and the political climate, to some extent it was also built into the design of the event itself. The G1000 took the form of a Town Hall Meeting, which is a format designed to facilitate the pooling and sharing of ideas and perspectives on a problem. As such, the proposals that were launched at the G1000 were still very open-ended, and not fit for immediate implementation.

The organizers therefore planned an additional phase in the project, which is called the G32. This G32 aimed to elaborate readily implementable proposals out of the basic ideas that were launched at the G1000. The G32 took the format of a Consensus Conference, which is often used in policy preparation processes throughout the world (Hendriks, 2005). Such a deliberative design is much more intensive and requires much more skill from the moderators in order to enable participants to propose specific policies and actions. It is also very open since the participants have a great say in the process itself. In fact, the citizens have the lead on what precisely they wish to work on and on how they want to work.

Nevertheless, the proposals of the mini-public have hitherto not been able to set the political agenda or influence parliamentary debates. Our analysis of parliamentary debates indicates that 12 references were made to the project since it ended. Only two of them mentioned the G1000 proposal on the age of retirement; the others dealt more with the process than with the results of the G1000. Moreover, some politicians stated that they will read through the text and that the ideas will certainly be kept in mind for future policies, but the results of the G1000 are likely to have little impact on the political agenda. As such, the weight given to the results is small.

Nonetheless, the process itself has already been replicated at smaller scales throughout the country and in neighboring countries with names directly inspired

by the G1000, such as the G100 or C5000. Hence, while the weight of the results in terms of outputs is low, the weight of the process is quite high.

Responsiveness and accountability

As the results from the G1000 and the G32 were not taken up in the policy-making process, there is no way of measuring the responsiveness of the implementation and the accountability of policymakers. The official responses of the political elites to the work of the citizens' panel were polite but they did not entail any political commitment, nor did the political elites feel that they should be held accountable in any way. As such, the G1000 scores badly when it comes to holding the elites accountable.

With regard to responsiveness, the picture is somewhat more nuanced. We organized a survey among a random sample of the Belgian population, which has shown that the results and proposals from the G1000 did tap onto wider public opinions. The respondents were asked to evaluate six of the proposals and all but one, which dealt with a universal basic income, found a large majority to support them. In this sense, the mini-public's opinions did find some echo in the wider maxi-public.

Conclusion

We started this article with the general claim that deliberation can only live up to its ambitions of improving the legitimacy of the system by respecting the principles of legitimacy in its own practices. We therefore set out to assess the democratic legitimacy of the G1000, a very specific deliberative event that took place in Belgium in 2011–2012. The G1000 was one of the very few deliberative processes that were organized in a truly bottom-up manner, and this made it a particularly interesting case for testing the potential of deliberation for generating democratic legitimacy.

Three findings are noteworthy. First of all, the G1000 did not do that well on all dimensions of legitimacy. As Table 1 shows, it scored very highly on the input dimensions. The quality of representation was good and the agenda could not have been more open. At the other end, the output legitimacy suffered from serious flaws. So far, no political weight whatsoever was given to the results and no feedback was set-up to keep the participants up to date about the implementation. With regard to the throughput legitimacy, the G1000 shows mixed results. The quality of participation was good thanks to a clear script and trained moderators, but the processes of aggregation at the G1000 were insufficiently transparent. Moreover, the political and media spin on the G1000 might have influenced the process for the worse.

The second finding is that the internal legitimacy of the G1000 to some extent depended on its design. As the project was crowd funded, the organizers had the opportunity to start with a very open agenda, which immediately boosted the project's input legitimacy. However, the absence of formal links to the main political

Table 1: Assessment of the internal legitimacy of the G1000

Dimension	Measurement	Arguments	Assessment
Input	Quality of representation	● Randomization ● Strong self-selection and dropout	Positive
	Openness of the agenda	● Crowd funding ● Clustering not transparent	Positive
Throughput	Quality of decision making	● Aggregation and deliberation ● Clustering central desk not transparent	Neutral
	Quality of participation	● Clear script with changing interaction styles ● Trained moderators	Positive
	Contextual independence	● Media and political spin	Negative
Output	Public endorsement	● No formalized public vote on the results ● Connection to wider public through G'Home and G'Off's	Neutral
	Weight of the results	● No implementation ● Low political commitment	Very negative
	Responsiveness and accountability	● No feedback loops	Very negative

actors meant that the organizers could not guarantee any formal implementation of the results. Hence, the design characteristics that increase input legitimacy also undermine output legitimacy.

Our final finding is that we have discovered the contours of a deliberative democratic dilemma. It is very difficult for a democratic project to be high on input, throughput and output legitimacy at the same time. In the case of the G1000, we clearly see that the organizers were able to include a wide variety of voices and to open up the agenda simply because they presented themselves as an independent political project. This independence (in terms of funding and organization) might very well have increased the input legitimacy, but it also undermined the output legitimacy. As the official political actors were mere observers on the sideline without any inputs into the process, the agenda or the funding, the G1000 completely lacked any binding commitments from the official instances when it came to implementing the results.

Moreover, we also discovered a trade-off between input and throughput legitimacy. When the quality of representation is high, the group is very diverse and the conflicts will be deep. This has implications for the throughput phase, because a diverse group is harder to manage from a group dynamic point of view. A high input legitimacy therefore does not necessarily induce a high throughput legitimacy.

The democratic trade-off between input, throughput and output legitimacy thus also threatens deliberative democracy and not only representative democracies. Deliberative democracy, which often presents itself as a way to improve both citizen involvement and participation in the input and throughput phases, and the quality of decisions at the output side, seems to suffer from the same problems as the types of

democracy it wants to be an alternative to. To see whether there is in fact a trade-off between the different dimensions of deliberation, a comparison between different projects with different designs could prove an interesting venue for future research. Moreover, such a comparison is all the more interesting and feasible given the recent increase in deliberative practice.

About the Authors

Didier Caluwaerts is a Democracy Fellow at the Harvard Ash Center, and a post-doctoral fellow of the Research Foundation – Flanders (FWO) at the Department of Political Science of the Vrije Universiteit Brussel. His research focuses on deliberative democracy, democratic innovations and experimental methods. He has recently published *Confrontation and Communication. Deliberative Democracy in Divided Belgium* (Peter Lang Publishers), which received the 2012 ECPR Jean Blondel PhD Prize.

Min Reuchamps (PhD) is professor of political science at the Université catholique de Louvain, Belgium. His research interests are federalism and multilevel governance, political sociology and democratic innovations. He has recently published *Les systèmes électoraux de la Belgique* (with Frédéric Bouhon, Bruylant) and *Les relations communautaires en Belgique* (with Julien Perrez, Academia-L'Hartmattan). He is the co-editor of a special issue on *The future of Belgian federalism* (with Kris Deschouwer, Regional and Federal Studies).

References

Bekkers, V. and Edwards, A. (2007) Legitimacy and democracy: A conceptual framework for assessing governance practices. In: V. Bekkers, G. Dijkstra, A. Edwards and M. Fenger (eds.) *Governance and the Democratic Deficit: Assessing the Legitimacy of Governance Practices*. Aldershot, UK: Ashgate, pp. 35–60.

Bohman, J. (2007) Political communication and the epistemic value of diversity: Deliberation and legitimation in media societies. *Communication Theory* 17(4): 348–355.

Caluwaerts, D. (2011) Van stemmen naar praten: Het ideaal van de deliberatieve democratie. *Samenleving en Politiek* 11(9): 79–85.

Caluwaerts, D. and Deschouwer, K. (2013) Building bridges across political divides: Experiments on deliberative democracy in deeply divided Belgium. *European Political Science Review*, published online, doi:10.1017/S1755773913000179.

Caluwaerts, D. and Reuchamps, M. (2012) The G1000. Facts, figures and some lessons from an experience of deliberative democracy in Belgium. In: Re-Bel (ed.) *The G1000*. Brussels, Belgium: Rethinking Belgium.

Caluwaerts, D. and Reuchamps, M. (2014) Does intergroup deliberation foster intergroup appreciation? Evidence from two deliberative experiments in Belgium. *Politics*, published online, doi:10.1111/1467-9256.12043.

Caluwaerts, D. and Ugarriza, J.E. (2012) Favorable conditions to epistemic validity in deliberative experiments: A methodological assessment. *Journal of Public Deliberation* 8(1): 1–20.

Cohen, J. (1997) Deliberation and democratic legitimacy. In: J. Bohman and W. Regh (eds.) *Deliberative Democracy: Essays on Reason and Politics*. Oxford: Blackwell, pp. 87–106.

Deschouwer, K. and Reuchamps, M. (2013) The Belgian federation at a crossroad. *Regional & Federal Studies* 23(3): 261–270.

Dryzek, J.S. (2000) *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford: Oxford University Press.

Dryzek, J.S. (2001) Legitimacy and economy in deliberative democracy. *Political Theory* 29(5): 651–669.

Easton, D. (1965) *A Systems Analysis of Political Life*. New York: Wiley.

Edwards, A. (2007) Embedding deliberative democracy: Local environmental forums in The Netherlands and the United States. In: V. Bekkers, G. Dijkstra, A. Edwards and M. Fenger (eds.) *Governance and the Democratic Deficit: Assessing the Legitimacy of Governance Practices*. Aldershot, UK: Ashgate, pp. 165–182.

Elliott, J., Heesterbeek, S., Lukensmeyer, C.J. and Slocum, N. (eds.) (2005) *Participatory Methods Toolkit: A Practitioners Manual*. Brussels, Belgium: King Baudouin Foundation.

Fishkin, J.S., Luskin, R.C. and Jowell, R. (2000) Deliberative polling and public consultation. *Parliamentary Affairs* 53(4): 657–666.

Fournier, P., van der Kolk, H., Carty, R.K., Blais, A. and Rose, J. (2011) *When Citizens Decide: Lessons from Citizens' Assemblies on Electoral Reform. Comparative Politics*. Oxford: Oxford University Press.

G1000. (2012) *Final report: Democratic innovation in practice*. Brussels, Belgium: G1000 : http://www.g1000.org/documents/G1000_EN_Website.pdf, accessed 23 September 2013.

Hendriks, C.M. (2005) Consensus conferences and plannning cells. In: J. Gastil and P. Levine (eds.) *The Deliberative Democracy Handbook*. San Francisco, CA: Jossey-Bass, pp. 37–58.

Hendriks, F. (2010) *Vital Democracy*. Oxford: Oxford University Press.

Huckfeldt, R. (2001) The social communication of political expertise. *American Journal of Political Science* 45(2): 425–438.

Krueger, R.A. (1998) *Moderating Focus Groups*. Thousand Oaks: Sage.

Landemore, H. (2012) Why the many are smarter than the few and why it matters. *Journal of Public Deliberation* 8(1): 1–12.

Manin, B. (1987) On legitimacy and political deliberation. *Political Theory* 15(3): 338–368.

Page, S.E. (2007) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton, NJ: Princeton University Press.

Papadopoulos, Y. and Warin, P. (2007) Are innovative, participatory and deliberative procedures in policy making democratic and effective? *European Journal of Political Research* 46(4): 445–472.

Parkinson, J. (2006) *Deliberating in the Real World: Problems of Legitimacy in Deliberative Democracy*. New York: Oxford University Press.

Reuchamps, M. (2011) Le G1000. *Politique: Revue des débats* 72: 64–66.

Reuchamps, M. and Caluwaerts, D. (2013a) Le tournant participatif et délibératif. In: F. Claisse, C. Laviolette, M. Reuchamps and C. Ruyters (eds.) *La participation en action*. Brussels, Belgium: P.I.E.-Peter Lang, pp. 9–21.

Reuchamps, M. and Caluwaerts, D. (2013b) En défense du G1000. *Politique: Revue de débats* 79: 49–51.

Rondas, J. (2011) De G1000 en het democratisch verdriet van België, http://opinie.deredactie.be/2011/11/26/de-g1000-en-het-democratisch-verdriet-van-belgie/, accessed 2 January 2013.

Ryfe, D.M. (2005) Does deliberative democracy work? *Annual Review of Political Science* 8(1): 49–71.

Saward, M. (ed.) (2000) Less than meets the eye: Democratic legitimacy and deliberative theory. In: *Democratic Innovation: Deliberation, Representation and Association*. London: ECPR Studies in European Political Science, pp. 66–77.

Schmidt, V.A. (2013) Democracy and legitimacy in the European Union revisited: Input, output and throughput. *Political Studies* 61(1): 2–22.

Szoc, E. and Lechat, B. (2012) La question cachée du G1000: Comment refaire société? *Emulations* 10: 41–47.

Smith, G. (2009) *Democratic Innovations: Designing Institutions for Citizen Participation*. Cambridge, UK: Cambridge University Press.

Surowiecki, J. (2004) *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York: Doubleday.

Thompson, D.F. (2008) Deliberative democratic theory and empirical political science. *Annual Review of Political Science* 11(1): 497–520.

van den Broeck, K. (2011a) David, koop een lidkaart! *De Morgen* 12 November 2011.

van den Broeck, K. (2011b) Burgertop is blaam voor parlement, http://www.knack.be/nieuws/burgertop-is-blaam-voor-parlement/article-opinion-20135.html, accessed 23 September 2013.