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

Min Reuchamps (Université catholique de Louvain)

Didier Caluwaerts (Vrije Universiteit Brussel)

First draft – do not quote without permission of the authors

From the beginning of the 1990's onwards, political analysts in all Western European countries discovered the contours of what they thought to be a widespread crisis of democracy. The alleged decline of political trust and public participation, 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, shape and function 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 (Dryzek 2000). A community of international scholars and philosophers, inspired by the work of Jürgen Habermas, became more and more convinced that a vibrant democracy is more than the aggregate of its individual citizens, and that democratic politics should be about more than merely voting. The quality of a democracy and 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. Democratic decisions were thus no longer considered a function of mere compliance with aggregation rules. Instead, they are determined by extensive argumentation about political choices before voting on them.

Because of its strong focus of public involvement in politics, this deliberative model of democracy started out in life as a theory of legitimacy (Benhabib 1996; Cohen 2002; Dryzek 2001; Parkinson 2006). By including everyone who is affected by a decision in the process leading to that decision, deliberation has important political merits: it is capable of generating political decisions that receive broad public support, even when there is strong disagreement on the aims and values a polity should promote (Geenens & Tinnevelt 2007, p. 47). After all, talking about political issues allows citizens to hear other perspectives to a problem and to see

their own perspectives represented in the final decision. As such, deliberative democracy seeks to score high on input, throughput and output legitimacy.

However, deliberation's 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 events 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.

In this paper, we set out to assess the internal legitimacy of one specific deliberative event, namely the G1000 project in Belgium (Caluwaerts & Reuchamps, 2012a). Our research question is therefore: to what extent does the G1000 live up to the criteria of input, throughput and output legitimacy? The G1000 project takes a particular place in the world of deliberative practice because it was not only grass roots in its process and its results, but also in its organization. Most deliberative events are introduced and funded by either public administrations or scientific institutions. The G1000 was rather considered 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. So instead of a scientific experiment, the G1000 was more of a democratic experiment by, through, and for citizens. This grass-root structure makes it a very interesting case for students of legitimacy, because as we will see later on it situated at the heart of the democratic trade-off between input and output legitimacy.

1   Measuring the internal legitimacy of deliberative events

In order to assess the quality of the deliberative process and its legitimacy, we rely on the traditional conceptualization of democratic legitimacy and distinguish between input, throughput and output legitimacy (Papadopoulos & Warin 2007; Risse & Kleine 2007). However, in the context of deliberation, these three dimensions of legitimacy have to meet slightly different criteria. In order to determine these criteria, we rely on Bekkers & Edwards' (2007) and Edwards' (2007) framework on deliberative legitimacy. They apply the concepts of legitimacy to the internal functioning of deliberative events, and distinguish between the dimensions taken up in table 1.

Table 1: dimensions of deliberative legitimacy

Input legitimacy	Quality of representation
	Openness of the agenda
Throughput legitimacy	Quality of decision making
	Quality of participation
	Contextual independence
Output	Weight of the results
	Responsiveness and accountability
	Preference transformation

1.1   Input legitimacy

The input legitimacy of deliberative events deals with citizens' opportunities to influence the agenda, the process and the outcomes of the deliberation. Input legitimacy is thus a measure for the openness of the deliberative events towards inputs, demands and needs from its participants. Input legitimacy is high if citizens have the chance to fully participate in the process of deliberation, and to make their viewpoints known. Inclusiveness is thus the central principle, whereas selectivity in demands and participants 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.

1. *Quality of representation*: A deliberative event can only be legitimate on the input side if its participants are in some way representative of the larger population from which they are drawn. Deliberation stands by the principle that a diversity of voices is a necessary requirement for good and legitimate decisions, and cognitive or experiential diversity should therefore be part of the mini-public (Thompson 2008).

    Forms of descriptive representation in deliberative fora through random sampling are therefore often considered to be ideal. They offer each member of the population an equal chance of being part of the mini-public. 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, even though these techniques are often used in designing deliberative democracy (Ryfe, 2005).

2. *Openness of the agenda*: Besides the fact that deliberative mini-publics have to be representative of the public as a whole, legitimate deliberation also requires an open agenda. Selectivity during the agenda setting of the event will undermine the input legitimacy because issues are banned from the discussion. Since policy problems are often holistic and interdependent, 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 could thus increase the input legitimacy.

1.2   Throughput legitimacy

Whereas the input phase mainly deals with the principle of inclusivity (of participants and their ideas), the throughput phase focuses on the deliberative process itself. The procedures used to steer the process of deliberation determine to a large extent the quality of that deliberation and its outcome. It is therefore crucial to critically assess the throughput legitimacy based on the following three criteria:

1. *Quality of decision-making*: First of all is it important that decisions in a deliberative setting come about through a process of argumentation, and not through coercion. Decisions must reflect the reasoned opinion of all those involved and not the power relations in the group. However, most deliberative designs do use 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 sheer number of times aggregation crosses the deliberative process and the binding power of these aggregative decisions. If for instance 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.

2. *Quality of participation*: It is not sufficient that diverse voices are represented in the deliberative arrangement; these voices also have to be heard. The deliberation has therefore to be modelled 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 towards the arguments of others. In this respect, the quality of

participation criterion captures whether the discussion reflects the characteristics of the ideal speech situation.

Crucial determinants of the quality of participation are therefore the style of moderation (active or passive involvement of the moderator), the setting (hot vs. cold deliberative settings), the size of the groups, the group dynamics and other inhibiting factors such as multilingualism and the issue polarization.

3. *Contextual independence*: Deliberative events never take place in a vacuum. They are always embedded in a certain political and societal context, and these contexts also exert influence on the process of deliberation. There can be large outside pressures on the participants to choose a particular course of action. In this sense, the throughput legitimacy is high when the deliberants are substantively independent from these outside pressures from political parties, public opinion makers, pressure groups or the media.

1.3   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:

1. *Weight of the results*: In order for the outputs to be legitimate, they have to be effectively implemented. This means that the outcomes of deliberation have to be in some way linked to formal political decision-making processes. Otherwise the process is merely a form of democratic *spielerei* with not practical use whatsoever. This means that the output legitimacy increases when the effective impact of the deliberative outputs and the participants on real world politics is high. It also means that the output legitimacy increases as the deliberative outcomes are binding: outcomes with mere advisory function score lower than when government has expressed its commitment to implementing the final decisions.

2. *Responsiveness and accountability*: Decisions taken through deliberation should not only be implemented by public institutions; there should also be regular feedback to the participants. Those who put their heads together to come up with solutions, should be kept in the loop on what happens with these solutions, what changes were made, and what problems the government agencies encountered. Moreover, there should be a clear chain of responsibility. It should be clear to the participants in deliberative events

who is going to implement their decisions, who is going to report on the progress made, and who can be held accountable for the results achieved.

3. *Preference transformation*: The outputs of deliberative events are not only situated in the wider public sphere. Deliberative democracy also assumes that there are important outcomes for the participants themselves. Preference transformation and the revision of previously held opinions is considered to be an important societal outcome of deliberation, and it should also be taken into account when assessing the output legitimacy of deliberative events. We therefore also have to look at the satisfaction among participants and opinion change as outputs.

2   The G1000 project

The G1000 is a large-scale deliberative project that took place in Belgium in 2011. Much in line with the political analyses that were made in the past two decades, the organizers of the G1000 project started from the idea that democracy is in crisis[1] (Caluwaerts, 2011). The organizers started from 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. Citizens are no longer satisfied with their role as passive subjects, and politicians are afraid to turn power over to the citizens. This public passivity is a particularly strong problem for the Belgian consociation, which relies solely on the elites for political decision-making.

Politicians and citizens are drifting further apart, whereas a modern take on citizenship and politics demands that they grow closer together. As a reaction to this widening gap, the G1000 presented itself as a citizen initiative that is capable of innovating democracy. Its aim was to gather ordinary citizens in a setting, which is conducive to open and uncoercive deliberation on contentious political issues, and to let citizens themselves experience the real world of democratic decision making with all its problems.

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. The agenda of the citizen summit itself was not determined by the organizers, as it is a commonplace practice in deliberative ventures. Rather, the organizers were convinced of the importance of starting with a very open agenda, which would be determined entirely by the public itself. In the beginning of July 2011, the

---

[1] The manifesto of the G1000 is available at on the website: http://www.g1000.org/.

organizers launched a so-called idea-box on the website in which every citizen, no matter what his opinion or background, could post questions or problems 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 through the media, citizens were invited to vote for their three preferred themes for the G1000. 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, following the model of a Town Hall Meeting. 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. The citizens' summit in Brussels was flanked by two side projects: G'Home and G'offs. The G'home was a software application allowing citizens to log into 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 via a citizens' panel. These 32 participants gathered during three weekends to reflect and write down proposals that are ready for implementation. 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.

3   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. However, in order to make claims about the impact of the G1000 on the Belgian political system, we should assess the legitimacy of the event 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.

3.1   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.

3.1.1   Quality of representation

The key issue for the input legitimacy of any deliberative event concerns the selection of the participants. Normatively and methodologically, the most appealing technique for recruiting participants of deliberative events is random selection (Bohman, 2007, pp. 351-352; Caluwaerts & Ugarriza, 2012; Fishkin & Farrar, 2005; Reuchamps, 2011). The reason why randomization is so normatively appealing is because it gives every citizen an equal chance of being selected to participate and leads to a high quality of representativeness of the mini-public. Moreover, 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, Luskin, & Jowell, 2000, p. 660). Only under these circumstances can deliberation generate 'better' decisions (Caluwaerts & Ugarriza, 2012)

This is also the reason why the G1000 opted for random selection. Besides methodological soundness, 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 and opinions (Caluwaerts & Ugarriza, 2012; Thompson 2008). When everyone at the table shares the same opinion, there is very little contestation within the group, and under such circumstances, deliberation does not lead to well-considered opinions and well-argued positions.

The participants were recruited through Random Digit Dialing. This technique generates random phone numbers for fixed and mobile lines and in Belgium has a penetration rate of 99%. Every inhabitant thus had an equal chance of being selected for participation in the G1000. In order to guard over the quality of the participant sample, the random selection was checked for certain predefined population quotas. Indeed, the selection guaranteed that

the sample resembled the population with regard to gender, age and province. This last quorum was considered crucial in order 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. 52% of the participants was female, 48% was 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% of the participants were Dutch-speaking versus 39% of the French speakers, which is also an accurate reflection of the population. And there was a large diversity in age groups, with the youngest participant being 18, and the oldest one being 85.

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 there were 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% of those invited eventually said yes. The commitment required for participating in a deliberative event thus proved to be very high since these citizens were asked to spend one day to discuss topics about which they often have no clue and possibly no interest. Moreover, because non-response rate was likely to be higher within minority groups, 10% of the seats were reserved for participants recruited not through random selection but through grassroots organizations engaged in helping non literate or with low socio-economic status people. So, while 90% of the participants were randomly selected, 10% came from a targeted recruitment. Such design aimed at maximizing inclusiveness.

In the end, the organizers did get a confirmation from 1000 participants, but they then experienced a dropout rate of about 30% among the people who had previously confirmed their participation. The final number of participants 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 socio-demographically 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% 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.

3.1.2   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 and the high levels of issue polarization in Belgian politics, 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 in the following phase put on the website and, once again through a media campaign, the public was asked to rate their top 3 issues. These issues would form the basis of the discussions at the citizens' summit.

The agenda setting process was thus very open, and the aim was to let the public agenda (not the political agenda) decide what was going to be discussed at the G1000. Indeed, it appeared that issues related the hottest political topic at the time, the so-called state reform, were not in the top 3, and they only figured marginally in the top 25. The G1000 thus relied entirely on the direct inputs from the citizens on which issues had to be discussed, which is why the project scores relatively 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 on the website (sic)" (G1000, 2011, pp. 5-6).

This high level of input legitimacy resulted to a large extent 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. This means that the organization did not depend on large external contributors, who would inevitably determine the agenda of the event in return for their contributions. 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% of the total budget" (G1000, 2011, p. 6). This ensured that the agenda was open and that the discussions were not shaped to fit the agenda of pressure groups or government.

3.2 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.

3.2.1 Quality of decision-making

The decision-making process throughout the event combined both deliberative and aggregative mechanisms. At multiple times during the event, the organizers asked concrete inputs from each of the tables. Sometimes this was a mere Post-it with feelings or key words, but most of the time the organizers asked for specific proposals or ideas, and more importantly also the arguments behind these proposals. Just like deliberative theorists argue, these arguments were considered essential for the deliberative quality of the process, and every argument that was aired during the discussions was noted down. As such, equal consideration was given to each participant's perspective, and 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, which was the key information processing office. 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, and these clusters were resubmitted to the individual vote of the participants at the end of each discussion round. Using voting equipment, the participants could individually express their opinions and preferences on each of the solutions that circulated in the discussions. Even though it was strictly speaking not necessary to have this final vote, it made the results of the discussions very tangible to the participants, and it also gave them the opportunity to see where they situated themselves in the larger group of participants.

The decision-making at the G1000 was thus characterized by an alternation between deliberation and aggregation. In a first phase, ideas, opinions, perspectives and arguments were collected at the tables, which were subsequently put to the vote at the end of each round. According to Bekkers and Edwards (2007), such a combination significantly improves the throughput legitimacy of the process because it reconciles 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 by the people at the central desk. After all, the central desk was 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. This could imply – but not automatically – that some genuinely innovative and original ideas, which only appeared once or twice, did not make it to the final vote. This is not ideal in terms of legitimacy because, in a truly democratic citizens' process, the aggregation should be done by the deliberants themselves.

The clustering by the central desk, although it is a common practice with such citizens' summits or town hall meetings (Elliott *et al*., 2005), negatively affected the throughput legitimacy because some (a very limited number, however) of the participants indicated after the event that they felt like their opinions were not taken seriously, and they were disappointed that their ideas were not put to the vote. There was thus a loss of information during the process, and 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, even though it only had to process the templates from the tables themselves.

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) were discussed. This left too little time for thorough argumentation and information exchange. The groups could therefore not dig very deeply in to the issues under discussion, and the process of social learning which is crucial to deliberation was somewhat hindered. The time restrictions thus further constrained the quality of the decision-making. Nonetheless, it was, to some extent, a deliberate decision by the organizers who wanted the participants to discuss not only one topic but three in order to offer several propositions which could then be dig further during the third phase of policy preparation. Designing deliberative democracy automatically implies trade-off and decision upon them. It is very clear when one assesses throughput legitimacy.

3.2.2   Quality of participation

As we have seen before, the quality of representation in the G1000 project was high. However, this does not 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 the crowd".

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 received an intensive training beforehand. 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 (Caluwaerts, 2012). 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 (Krueger, 1998). A good start is an essential prerequisite for high quality participation.

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. In 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. They were too one-sided and mostly offered rather left-wing views on the issues at stake. This 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, these negative findings were 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 to lower the threshold of speaking in public; at other times, they discussed the issues with everyone at the table. Sometimes the tasks were very simple like expressing their feelings towards an issue such as immigration, but at other times, the expected outcomes were more complex and the cognitive investment was much larger. 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 made the discussions very effective in bringing out the cognitive diversity of the group.

3.2.3   Contextual independence

Because of its timing, the G1000 was embedded in a very specific social and political context. The event took place at a moment when the negotiations between the party leaders of the north and the south of the country were completely deadlocked. Bringing citizens together in such a controversial political climate has consequences for the 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 attempting to delegitimize it in Flanders. 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.

The context thus mattered, but because the agenda setting was open, the topics to be discussed were not imposed by any political agenda. In fact, the three issues that received

most votes were not – directly – related to the so-called community problems between Flemings and Walloons. Of course, the participants could, if they wished, bring in topics related to the state reform while discussion the three main topics (social security, welfare in times of economic crisis and immigration). Furthermore, the final round of discussion was left open to the choice of each table. Out of the 81 tables, only 14 chose to discuss an issue related to the governing of divided Belgium, all the others tackled topics such as energy policy (16), links between school and job market (12), renewal of democracy (7), mobility (6) and innovation and creativity as boosters for our economy (5).

In sum, while the project itself was much influenced by the political context and thus its independent from the political and social environment was rather limited, the deliberative design and the participants themselves could step aside from this context of political turmoil.

3.3   Output

After decisions are made – whether through deliberation or aggregation – these decisions have to be implemented before it can generate any lasting outcomes. In this section, we look at what became of the results of the G1000, and how the political elites were held accountable. Moreover, we look at whether the participation in the G1000 resorted any effects on the participants themselves.

3.3.1   Weight of the results

Decisions generated through deliberation can only be legitimate on the output side when they are effectively implemented. 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 so far – at least after the first two phases of the project. The G1000 grew as a truly bottom-up initiative, in the sense that it was initiated and funded entirely by citizens. This had the advantage that the event had a very open agenda (as we saw earlier), but the flipside of the coin is that there was no commitment whatsoever from the political elites to implement the results from the G1000. It was looked at with great scepticism by political parties, and apart from the presidents of the different Belgian parliaments saying that it was important to listen to the citizens, very little specific action was taken by politicians to implement the ideas of the G1000.

However, the lack of implementation of the results was not only due to the bottom-up organization and funding of the G1000, to some extent it was also built into the design of the event itself. The G1000 took the format of a Town Hall Meeting, as mentioned earlier. Town Hall Meetings are well known formulas in deliberative practice. They are large events 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 concrete proposals the basic ideas that were launched at the G1000. The G32 took the format of a citizens' panel or also called Consensus Conference, which is often used in policy processes throughout the world. 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 more open than a citizen summit since the participants have a much greater say in the process itself. In fact, the citizens do have the lead on what precisely they wish to work on (i.e. the choice of the specific questions they want to tackle), on how they want to work (i.e. the choice of the experts and stakeholders they wish to question) and above all on what they decide and bring to the public debate.

3.3.2   Responsiveness and accountability

Since the results from the G1000 were not yet taken up in the policy making process, there is no way of measuring the responsiveness of the implementation, and the feedback loops necessary for the accountability of policy makers. This aspect of the output legitimacy will have to be assessed after the G32, on the short-term, mid-term and long-term.

3.3.3   Preference transformation

Citizens can change the social and political infrastructures directly through making decisions with real political impact, but also indirectly by changing their opinions and bringing about political change in the long run. As in so many deliberative events, the participants in the G1000 experienced transformative effects through talking to each other. The comparison between the results of pre-questionnaire and the post-questionnaire show that they became more trusting of political elites and institutions, but at the same time they became more

supportive of citizen involvement in politics (Caluwaerts & Reuchamps, 2012a). They also became more positive towards the other linguistic group, especially when they were seated at bilingual tables. Contrary to what could have been expected in a context of deep community tension, as reminded in the discussion about the context, the opinions toward the other group (the outgroup) of the participants seated in a bilingual table (i.e. in a table where both Dutch-speakers and French-speakers were seated and where simultaneous translation was provided, so bilingualism was not required at all) were stable before and after the event, while the opinions of those who were seated in unilingual tables became more negative (Caluwaerts & Reuchamps, 2012b). In other words, being seated with participants of the outgroup acted as buffer against depreciation of the outgroup. Although it was not its aim, the G1000 thus showed that bringing together citizens from both sides of the linguistic border did neither increase the divide between their two communities nor fade away the sources of the conflicts but rather, and at least, could help establish a dialogue. This is not uninteresting regarding input, but especially throughput and output legitimacy.

4   Conclusion

We started this paper with the general claim that deliberation can only live up to its political 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. The G1000 was one of the very few deliberative processes that were organized in a truly bottom-up manner. It was funded by ordinary citizens, it was organized as a citizens initiative, and the agenda was set by the general public. This made it a particularly interesting case for testing the potential of deliberative events for generating democratic legitimacy, and three findings are noteworthy.

First of all, the G1000 did not do that well on all dimensions of legitimacy. As table 2 shows, it scored very highly on the input dimensions. The quality of representation was very good and the agenda could not have been more open. At the other end, the output legitimacy suffered from serious flaws. So far (but the third phase is not yet fully completed), no political weight whatsoever was given to the results and no feedback loops were 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, but the processes of aggregation at the G1000 were insufficiently transparent. Moreover the

reporting on the G1000 in the media and the opinions on the event by politicians will probably have influenced the process – for the worse.

Table 2: Assessment of the internal legitimacy of the G1000

Dimension	Measurement	Arguments	Assessment
Input	Quality of representation	• Randomization • Strong self-selection and dropout	++
	Openness of the agenda	• Crowd funding • Clustering not transparent	+++
Throughput	Quality of decision making	• Aggregation AND deliberation • Clustering central desk not transparent	+
	Quality of participation	• Clear script with changing interaction styles	++
	Contextual independence	• Media and political spin	-
Output	Weight of the results	• No implementation	---
	Responsiveness and accountability	• No feedback loops	---
	Preference transformation	• Clear process of opinion change	++

The second finding is that the internal legitimacy of the G1000 to some extent depended on the specific origins and design of the event. For instance, because the project was paid for using crowd funding, the organizers had the opportunity to start with a very open agenda, which immediately boosted the project's input legitimacy. Similarly, the fact that the project was organized by a citizens movement without any formal links to the main political actors (parties, social movements, etc.) or to the main political fora, meant that the organizers could not guarantee any formal implementation of the results. No weight was so-far given to the outcomes of the deliberation, which undermines the output legitimacy of the project.

A 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 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 completely open up the agenda simply because they presented themselves as an independent political project that would not be manipulated by the political elites. However, this strong independence (in terms of funding and organization) also came at a price. It might very well have increased the input legitimacy, but it also undermined the output legitimacy. Since 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 comes to implementing the results.

Input and output legitimacy are thus in a trade-off relationship, which is something Robert Dahl (1991) already hinted at over twenty years ago. The novelty of the present paper lies, however, in the fact that it shows that trade-offs also threaten deliberative democracy and not only representative democracies. Deliberative democracy, which often presents itself as a way to improve both citizen involvement at the input side, 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. And such a comparison is all the more interesting and feasible given the recent increase in deliberative practice.

References

Bekkers, V. & A. Edwards (2007), 'Legitimacy and Democracy: A Conceptual Framework for Assessing Governance Practices', in V. Bekkers, G. Dijkstra, A. Edwards & M. Fenger (eds.), *Governance and the Democratic Deficit. Assessing the Legitimacy of Governance Practices*. Aldershot: Ashgate, pp. 35-60.

Benhabib, S. (1996), 'Toward a Deliberative Model of Democratic Legitimacy', in Benhabib, S. (ed.), *Democracy and Difference*. Princeton: Princeton University Press, pp. 67-94.

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', in *Samenleving en Politiek* 18 (9), pp. 79-85.

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

Caluwaerts, D. (2012), 'Waar en wanneer spreken mannen en vrouwen over politiek? De sekseverschillen in politieke discussie in hun sociale en politieke context', in *Res Publica* 53 (1), pp. 5-26.

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

Caluwaerts, D. & M. Reuchamps (2012b), "Does intergroup deliberation foster intergroup appreciation? Evidence from two deliberative experiments in deeply divided Belgium". Paper presented at the bi-annual IPSA Conference, Madrid, July 2012.

Cohen, J. (2002), 'Deliberation and Democratic Legitimacy', in Estlund, D. (ed.), *Democracy*. Oxford: Blackwell, pp. 87-106.

Dahl, R. A. (1991), *Democracy and its Critics*. New Haven: Yale University Press.

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), pp. 651-669.

Edwards, A. (2007), 'Embedding Deliberative Democracy: Local Environmental Forums in The Netherlands and the United States', in in V. Bekkers, G. Dijkstra, A. Edwards & M. Fenger (eds.), *Governance and the Democratic Deficit. Assessing the Legitimacy of Governance Practices*. Aldershot: Ashgate, pp. 165-182.

Elliott, J., S. Heesterbeek, C. J. Lukensmeyer & N. Slocum (Eds.) (2005), *Participatory Methods Toolkit. A practitioners manual*. Brussels: King Baudouin Foundation and the Flemish Institute for Science and Technology Assessment (viWTA).

Fishkin, J. S., & Farrar, C. (2005). Deliberative Polling*: From Experiment to Community Resource. In J. Gastil & P. Levine (Eds.), *The Deliberative Democracy Handbook : Strategies for Effective Civic Engagement in the 21st Century* (pp. 68-79). San Francisco: Jossey-Bass.

Fishkin, J. S., Luskin, R. C., & Jowell, R. (2000). Deliberative polling and public consultation. *Parliamentary affairs, 53*(4), 657-666.

G1000 (2011), "Report by the international observers", 2011.

Geenens, R. & Tinnevelt, R. (2007). Het deliberative model van democratie. In R. Geenens & R. Tinnevelt (eds.), *De stem van het volk. Democratie als gesprek* (pp. 15-63). Leuven: Lannoo Campus.

Huckfeldt, R. (2001). The Social Communication of Political Expertise. *American Journal of Political Science, 45*(2), 425-438.

Papadopoulos, Y. & P. Warin (2007), 'Are innovative, participatory and deliberative procedures in policy making democratic and effective?', in *European Journal of Political Research* 46, pp. 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.

Risse, T. & M. Kleine (2007), 'Assessing the Legitimacy of the EU's Treaty Revision Methods', in *Journal of Common Market Studies* 45 (1), pp. 69-80.

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

Thompson, D.F. (2008), 'Deliberative Democratic Theory and Empirical Political Science', in *Annual Review of Political Science* 11 (1), pp. 497-520.