Crowdsourced Deliberation: The Case of the Law on Off-Road Traffic in Finland

Tanja Aitamurto and Hélène Landemore

This article examines the emergence of democratic deliberation in a crowdsourced law reform process. The empirical context of the study is a crowdsourced legislative reform in Finland, initiated by the Finnish government. The findings suggest that online exchanges in the crowdsourced process qualify as democratic deliberation according to the classical definition. We introduce the term “crowdsourced deliberation” to mean an open, asynchronous, depersonalized, and distributed kind of online deliberation occurring among self-selected participants in the context of an attempt by government or another organization to open up the policymaking or lawmaking process. The article helps to characterize the nature of crowdsourced policymaking and to understand its possibilities as a practice for implementing open government principles. We aim to make a contribution to the literature on crowdsourcing in policymaking, participatory and deliberative democracy and, specifically, the newly emerging subfield in deliberative democracy that focuses on “deliberative systems.”

KEY WORDS: crowdsourcing, crowd law, deliberative systems, online deliberation, participatory democracy, deliberative democracy, policymaking, democratic innovations

Introduction

Crowdsourcing has become a popular method for gathering knowledge in realms ranging from new product design to social science research (Aitamurto, 2012; Aitamurto, Holland, & Hussain, 2015; Smith, Richards, & Gastil, 2015). Crowdsourcing for policymaking in government, and particularly in lawmaking, is still, by contrast, a relatively new phenomenon (Aitamurto & Landemore, 2015; Christensen, Karjalainen, & Nurminen, 2015; Lehdonvirta & Bright, 2015). One of its best known instances occurred in Iceland in 2011, when the country used crowdsourcing in its constitution-writing process by allowing citizens to comment on 12 successive constitutional drafts published online (Landemore, 2015). Beyond the paradigmatic Icelandic experiment, there are a multitude of instances of crowdsourcing in both local and national governance across the world, as governments implement crowdsourcing as part of their open government practices aimed at fostering civic engagement and knowledge discovery for policies.
While crowdsourcing offers a promising alternative to traditional, more closed lawmaking and policymaking processes, it also raises a number of questions. First, is crowdsourcing conducive to deliberation among citizens or is it essentially a mere consulting mechanism for information gathering? Second, if it is conducive to deliberation, what kind of deliberation is it? Is it, in particular, democratic? Third, to the extent that democratic deliberation specifically requires inclusiveness of viewpoints and interests, how representative are the online deliberative exchanges of the wishes and priorities of the larger population?

Our case study is a crowdsourced law reform in Finland in which citizens were invited to contribute knowledge to law reform online. In this article we ask whether crowdsourcing—in this case for ideas and knowledge—allows for democratic deliberation among the participants. We also consider the implications for crowdsourced policymaking, the possibility of massive online deliberation, and participatory democracy.

We find that despite the lack of clear incentives for deliberation in the crowdsourced process, crowdsourcing functioned as a space for democratic deliberation, namely an exchange of arguments among participants characterized by a degree of freedom, equality, and inclusiveness. An important finding, in particular, is that despite the lack of statistical representativeness among the participants, the deliberative exchanges reflected a diversity of viewpoints and opinions, tempering to a degree the worry about the bias likely introduced by the self-selected nature of citizen participation. Lack of statistical representativeness thus does not necessarily mean a poverty of views, information, and arguments and low quality deliberation. We introduce the term “crowdsourced deliberation” to mean the deliberation that happens in crowdsourcing, even when the crowdsourcing process primarily aims to gather knowledge rather than to generate deliberation. Crowdsourced deliberation can thus be an intentional or unintentional product in a crowdsourcing process. We expand on this definition later in the article. All in all, the deliberation taking place in the Finnish experiment can thus be characterized as a democratic type of crowdsourced deliberation, that is, “crowdsourced democratic deliberation.”

The article is structured as follows. We first introduce the key concepts of democratic deliberation and crowdsourcing. We then introduce the case profile, data, and methods. The next section presents the findings about the deliberative and democratic nature of the online exchanges taking place on the platform and addresses the objection of the lack of statistical representativeness of the participants. We then conclude and sketch out further avenues of research.

**Key Concepts**

*Democratic Deliberation*

At an abstract level and as defined by so-called “classical” deliberative democrats (Mansbridge et al., 2010), democratic deliberation is “the public use of arguments and reasoning among free and equal individuals” (adapted from Cohen,
The “use of arguments and reasoning” can be further specified as an exchange of arguments in which the participants aim to convince their interlocutors of the validity of a claim or, conversely, to refute a given claim. Deliberation in that sense is distinct from bargaining, which consists in appealing to self-interest, or threatening (Elster, 1986). Democratic deliberation is also here meant as an intersubjective exercise among at least two individuals, as opposed to an internal dialogue in the vein of “deliberation within” (Goodin, 2005) or a deliberation occurring among entities larger than individuals, as in system-thinking.

Though deliberation is meant to be “rational,” nothing in the definition we endorse requires that the arguments be exchanged as explicit syllogisms. On the contrary we assume that they can be phrased more elusively, taking for example the form of stories or a series of anecdotes, which can be reconstructed or identified as arguments for or against something. Deliberation thus requires a reasoned exchange of arguments; democratic deliberation requires equal standing among free participants (“free and equal”) as well as the “public” aspect of the exchange.

More demanding conceptions of democratic deliberation emphasize other procedural aspects such as norms of mutual respect, reciprocity, and civility (as per Gutmann & Thompson, 2009), or truthfulness and sincerity, as well as the normative goal of a rational consensus (Habermas, 1996). Kies (2010, p. 42) thus lists a series of nine criteria that would need to be met in order for an exchange to qualify as “deliberative” according to the most stringent requirements of the literature on deliberative democracy. While, as we will point out, some of these criteria were at least partially met in the Finnish crowdsourcing experiment, we believe that even processes that fail to meet them all or even most of them can still play vital democratic and deliberative roles.

We thus set a more minimal bar for democratic deliberation than some deliberative democrats might be comfortable with. This is justified, in our view, by the fact that the design of our experiment was not aimed at generating the ideal discursive exchange theorized by deliberative democrats. Relaxing the criteria and focusing instead of what we see as core features of democratic deliberation makes it possible to see deliberation happening under even less than favorable conditions. We do not set the bar so low, however, as to fall under the threshold of minimal deliberation (as, in our view, others do). We thus follow Delli Carpini, Cook, and Jacobs (2004) in seeking to retain a distinction between deliberation per se and any kind of political talk or communication—what they call “discursive participation” and “citizen engagement.” The threshold criterion is, as we shall now explain, the presence or absence of arguments and critical listening.

Speaking in concrete terms, we take the definition of deliberation per se, as public exchange of arguments, to translate into a continuum of practices. This continuum of practices would at best embody “discussion that involves judicious arguments” and “critical listening” as well as “a careful examination of a problem or issue, the identification of possible solutions, the establishment or reaffirmation of evaluative criteria, and the use of these criteria in identifying an optimal solution” (Gastil, 2000, p. 22). At the lower extreme—poor deliberation—the practices would still involve arguments and critical listening of some kind, albeit of a lower quality.
Fishkin (1995, p. 41, cited in Delli Carpini et al., 2004, p. 317) provides a useful characterization of lower quality deliberation as “incomplete” when “arguments offered by some participants go unanswered by others, when information that would be required to understand the force of a claim is absent, or when some citizens are unwilling to weigh some of the arguments in the debate.” Following Chambers (2003, p. 309, cited in Delli Carpini et al., 2004, p. 317), we may also operationalize deliberation as “debate and discussion aimed at producing reasonable, well-informed opinions, in which participants are willing to revise preferences in light of discussion, new information, and claims made by fellow participants.”

Crowdsourcing: Open Participation in Policymaking

Crowdsourcing is an open call for anyone to participate in an online task (Brabham, 2008, 2013; Estelles-Arolas & González-Ladrón-de-Guevara, 2012; Howe, 2008) by submitting information, knowledge, or talent. Unlike in outsourcing, in which a task is assigned to a specific agent, crowdsourcing has no target group defined ex ante. “The crowd” refers to the individuals that self-select from a larger pool of people—in theory anybody who has access to the Internet and is aware of the task.

Crowdsourcing can be either voluntary or paid. The latter provides financial incentives, the former does not. Crowdsourcing can further be divided according to the kind of tasks outsourced to the crowd. Typical categories are microtasking, idea generation, knowledge search, and argumentation (Aitamurto & Landemore, 2015). In crowdsourcing for microtasking, organizations outsource tasks that are sufficiently small and simple to be performed by anyone willing (Kittur, Chi, & Sùr, 2008). Crowdsourcing for idea generation is used by companies for instance through innovation intermediaries such as InnoCentive, or on companies’ own platforms like Dell’s IdeaStorm. Journalists use crowdsourcing for knowledge search on the online platforms of news outlets and on social media (Aitamurto, 2015). Similarly, emergency management organizations use crowdsourcing to solicit information from volunteers in crisis situations (Liu, 2014; Starbird, 2011). Platforms like Consider.it (Kriplean, Morgan, Freelon, Borning, & Bennett, 2012) and Deliberatorium (Klein, 2011) facilitate crowdsourced argumentation and deliberation by allowing users to express a position in an online forum.

Crowdsourcing, as examined in this study, comes into play in the research and drafting stage—as opposed to the decision stage—of a legislative process in Finland, as illustrated in Figure 1. While the early stages of lawmaking involve crowdsourcing, it is the parliament that ultimately decides about the bill. As shown in Figure 1, civil servants in the government do research for the bill and draft it. They work directly with interest groups and with an expert committee that includes representatives from stakeholders in the policy. Once the government accepts the bill, it goes to Parliament, which decides on the law. The crowd brings in additional data points to the research and drafting part of policymaking. Though the crowd itself is largely anonymous, it can include representatives from interest groups, hence the overlap between the elements in Figure 1.
Case Profile, Methods, and Data

About the Off-Road Traffic Law

The case presented in this article is a partially crowdsourced off-road traffic law reform process in Finland. The off-road traffic law regulates traffic beyond established roads, that is, motor-powered transportation in the countryside, mainly with snowmobiles in the winter and ATVs in the summer. Off-road traffic is regulated by the Ministry of Environment under a law that came into effect in 1995.

There had been pressure in Finland to reform the law for several reasons, including the increased volume of off-road traffic. One of the previous governments of Finland (in power 2010–11) had proposed a bill to the Finnish Parliament to reform the off-road traffic law in 2010, but the bill expired in parliament after raising controversy. The Finnish Ministry of Environment with the Committee for the Future in the Finnish Parliament then decided to use crowdsourcing in the law reform process. As stated by the Minister of Environment, the goal was to search for knowledge and ideas from the crowd, enhance people’s understanding of the law, and attempt to increase the perception of the policy’s legitimacy. The authors of this article were in charge of implementing and analyzing the crowdsourcing experiment.

Crowdsourcing in Off-Road Traffic Law Reform

Crowdsourcing took place in two sequences in the spring of 2013 on an online platform (Figure 2). The participants could propose ideas on the platform,
vote others’ ideas up or down (Figure 2), and comment (Figure 3). The crowd-generated input was accessible online for both registered and non-registered visitors on the website. In order to leave a comment, propose an idea, or vote (thumbs up/down modality) on the platform, users had to register on the site. They could choose to stay anonymous, use their real names, or create a nickname. A verifiable email address was required for registration. In addition to the crowdsourcing platform, the Ministry set up a website\(^6\) to provide more access to information about the law.

The crowdsourcing process had two phases, as illustrated in Figure 2. Civil servants in the Ministry of Environment, who are experts on the law and who wrote the expired bill defined, together with the authors of this article, main areas for crowdsourced problem identification, including broad topics such as problems related to off-road traffic, and narrower ones, such as age limits for off-road traffic, emissions standards, and regulation of the route establishment process. The prompts for the participants included information about the law and questions for them to answer. Within each problem area, the participants could propose ideas and share their concerns and experiences about off-road traffic. There was also a category called “Propose your own topic,” which allowed the participants to make suggestions outside the provided framework.

The two crowdsourcing phases generated about 500 ideas and 4,000 comments, and 24,000 up or down votes from about 700 users. The researchers analyzed the participants’ input and, together with the civil servants in the Ministry, organized the ideas and comments into categories. Then the ideas were
assessed by crowd evaluation and expert evaluation (Lee, Goel, Aitamurto, & Landemore, 2014). While this evaluation phase was important in reporting the results to the Ministry of Environment in October 2013 (Aitamurto, Landemore, Lee, & Goel, 2014), we leave its analysis out of the scope of this article, because it was nondeliberative in nature. Since the law reform is not complete as of this writing,7 we also leave the analysis of the end-result, the reformed law, out of the article. We focus here on the question of democratic and deliberative aspects in the first two crowdsourcing stages.

Data and Methods

The authors participated in the design and planning of the crowdsourcing platform as advisors, thus applying the method of action research. In action research, the field is not something to be observed; rather, the researcher is active in interacting, producing, and creating the research site (Gustavsen, 2001). Once crowdsourcing began, the authors took the role of participant-observers (Hansen, Cottle, Negrine, & Newbold, 1998). The Finnish-speaking author was a moderator on the platform with two other moderators. The participation of the researchers helped to build a rapport with the interviewees and enabled a better understanding of the technical aspects shaping the process. The researchers’ participation did not hinder the interviewees from sharing negative experiences and expressing sincere critical opinions and observations. We used digital ethnography, interviews, and an online survey as data gathering methods.
Digital Ethnography. We used digital ethnography, also called netnography (Kozinets, 2002), to gather data from the online process. Ethnographic data collection began in January 2013, paused in April for the analysis between the two crowdsourcing phases, and continued from the end of April through the end of the online process on June 24, 2013. Daily observations (30 minutes to 1 hour per day for 153 days) were collected in field notes about the interactions on the online platform, and the notes were incorporated into memos. With this information we mapped the developments on the platform and turning points in the conversations, and observed the democratic and deliberative aspects of the interactions. The data informed our understanding of the crowdsourced process and shaped the design of the interview outline and survey.

Interviews With Key Informants. We interviewed—both on the phone and in person when possible—online participants, civil servants, and politicians involved in the crowdsourcing process. The interviewees were chosen based on their involvement in the process and their expertise. The interviewees were recruited by email through the online platform, and emails were sent to a random sample of the participants across the activity levels, from the most active to those who signed up but never commented or voted on any idea. However, those who responded positively to the interview request were all individuals who participated in an active manner in the crowdsourcing process. Thus, our sample excludes the most passive members of the crowd. The interviewees’ activity level (several ideas, comments, and votes) varies from high activity to low activity (no ideas, just comments, and votes).

A total of 21 online participants were interviewed, eight of whom were interviewed twice across the two phases. Six of the interviewed participants were female and 15 were men. The average age was 53 years, and the range was from 27 to 69 years. Seven of the interviewees were retired, and 12 were working in a range of occupations. They included an electrical engineer, a business and product manager, a kindergarten teacher, a lawyer, a wilderness guide, an environmental and land-use expert in municipal government, and a forest expert. The common denominator for the participants was that they had a reason to care about off-road traffic and the law regulating it.

We also interviewed the Minister of Environment, the Vice President of the Committee for the Future (the head of the crowdsourcing project and a Member of the Parliament), two civil servant experts in off-road traffic law in the Ministry of Environment, and a communication expert involved in the crowdsourcing process. Finally, we interviewed twice three members of two interest groups (The Central Union of Agricultural Producers and Forest Owners and the Finnish Association for Nature Conservation) involved in off-road traffic matters. In total, we conducted 40 interviews with 29 individuals. The average length of the interviews was 57 minutes. We used a semi-structured interview outline with questions focusing on perceptions about democracy, the experience of participation, and expectations for the outcome. The interviews were recorded and
transcribed. The interviewees are identified by numbers 1–25 when quoted in the text.

**Online Survey.** An online survey was designed to examine the participants’ demography and their perceptions and expectations about crowdsourcing in democracy, learning in crowdsourcing, the impact of crowdsourcing on law reform, and opinion change. The survey link was emailed to all participants on the crowdsourcing platform. Of 748 registered users, 204 started the survey and 186 completed it, resulting in a 25 percent response rate.8

We analyzed the interview data following Strauss and Corbin’s (1998) analytical coding system. In the first round, we used open coding, allowing key themes and patterns to emerge from the data and thus to guide further analysis, following the principles of grounded theory (Lindlof & Taylor, 2002, p. 214; Strauss & Corbin, 1998, p. 101). Coding involved dissecting each transcript paragraph by paragraph to identify recurring categories and themes. In the next coding round, we used axial coding to relate the emerging categories to subcategories (Saldana, 2009, p. 159; Strauss & Corbin, 1998, p. 123). The coding was conducted by three researchers, who coded the same passages and cross-checked the codes, finding agreement between different systems. Finally, we applied selective coding (Strauss & Corbin, 1998, p. 143) to integrate and synthesize the subcategories into main categories: democratic aspects and deliberative aspects in crowdsourced policymaking, referring to how the interviewees perceive democracy, representativeness, and deliberation in crowdsourcing.

The ethnographic data were analyzed using these categories as a framework. We used the memos to find examples of democratic deliberation as we define it (as characterized by exchange of arguments, including in the form of personal stories, and by critical listening among the involved parties). The excerpts from the ethnographic data in the Findings section below are chosen to show typical examples of deliberation on the platform.

**Findings**

*Democratic Deliberation in Crowdsourcing*

To what extent did the Finnish experiment, which primarily aimed to be an idea and knowledge search and was not designed to encourage deliberation, nonetheless produce the kind of deliberation deliberative democrats are interested in? What we found is that participation in crowdsourcing generated democratic deliberation in the following ways.

First, the participants exchanged arguments with others in the dialogical and intersubjective manner typical of the ideal of democratic deliberation described by deliberative democrats. The exchanges also display, to a varying degree, elements of Gastil’s (2000) and Chambers’ (2014) more concrete definitions, which
we will point out as relevant. We give below two typical representative examples of deliberative online exchanges.

*Example 1.* Our first example is an exchange on the crowdsourcing platform about the appropriate age limit to ride snowmobiles and tractors. In a thread (started January 23, 2013), a participant argued that age limits for off-road vehicles were unnecessary, and in fact, are counterproductive as learning from a young age in nature and under parental supervision is what makes for experienced and safe drivers. The participant gave their own story as an example in support of the argument:

Commenter A: “I started driving a tractor on forest roads and fields when I was six years old. My six children started driving mopeds when they were five years old and snowmobiles when they were seven years old. None of them has ever had any accidents and I’m now in retirement!”

Three supporting views were presented, after which came the counterpoint that the most dangerous drivers are often 16-year-old boys. The person argued that age limits should apply to snowmobiles specifically because of their similarities with motorcycles (please note our emphasis on the appeal to reasons):

Commenter B: “Snowmobile is probably the fastest vehicle that can legally be driven by a 15-year old. It is not that different from a motorcycle. Driving fast with a snowmobile also requires physical strength, and not all 15-year olds have that. *Because of the aforementioned reasons,* if I was the one deciding, for riding a snowmobile on a route one should be 18 years old. The younger ones should be able to ride only on closed tracks, or, on private property.”

This point was then reinforced by two more comments before a third participant introduced yet another ripple in the argument. After acknowledging that different age limits should be considered for different types of off-road traffic vehicles, Commenter C argues that regardless of what the law says:

People should use common sense. Just because the law allows it does not mean that parents should let their sons ride freely the fastest snowmobiles. In the countryside, instead, people have always been driving and will be driving a tractor from an early age on. […] And I don’t see that there’s anything wrong with that as long as driving happens under monitored circumstances until a certain age.

A fourth commenter then introduced a distinction between recreational and professional use of off-road traffic vehicles, suggesting that age limits should be increased for recreational but not professional use (please note our emphasis on
Commenter D: “On the tracks and routes the age limit for snowmobile riding could be considerably increased, compared to motorcycles, because a snowmobile is a very fast vehicle and the youngsters often tend to ride around habitat centers and small ice-covered areas, thus causing unnecessary disturbance to others. A 15–17 year old doesn’t have an understanding of the responsibilities related to riding such a vehicle yet. In particular they don’t understand that they cause unnecessary disturbance.”

This comment was followed by an opposing comment supporting lower age limits without much argument, before a fifth argument points out an alleged reasoning flaw in the underlying logic of the existing law in its focus on facilitating farm work as opposed to minimizing the probability of severe accidents (please note our emphasis on the mention of reasoning flaws):

Commenter E: “In the new (expired) bill the lowering of the age limits is broadly rationalized as necessary because of the participation of underaged children in farm work. This reasoning puts the issue on the wrong track. The problems of age limits are related to large accident risks, which the actual off-road traffic vehicles cause, in part because they are getting faster and faster. A tractor used for farming is not an off-road traffic vehicle comparable to a snowmobile or an ATV.”

This thread displays at least five distinct arguments that answer each other specifically and complicate the debate in productive ways. It starts with a story about how one of the participants taught their own children to ride when they were under 10 years old, suggesting that children younger than 15 are perfectly able to ride snowmobiles, at least when properly supervised by adults. It continues with opposing viewpoints that bring up counter-examples (the speed-crazed 16-year-old snowmobile riders), counter-arguments, and conceptual distinctions (between types of off-road traffic vehicles and recreational and professional use), all of which result in new proposals for custom-tailored age limits per category of off-road vehicles.

Clearly, the discussion involves decent enough arguments, some of which arguably count as “judicious” (as per Gastil’s [2000] requirement). It also displays a certain level of “careful examination of a problem or issue” as well as “the identification of possible solutions” (again, as per Gastil’s [2000] definition). Additionally, the tone of the exchange is respectful and individuals display “critical listening,” taking opposing views’ seriously.

Example 2. Another thread (Example 2, a thread started on the 20th of June, 2013) involved proponents of the need-based and rights-based approaches to the
creation of new routes. This exchange illustrates how "evaluative criteria" are posited and used to identifying optimal solutions (as per Gastil, 2000, p. 22). Proponents of the need-based approach argue that routes should be created only if there is a demonstrated need for them, whereas advocates of the rights-based approach argue that snowmobilers have a basic right to freedom of movement. 

A number of arguments, appealing to personal experience or more abstract libertarian principles, were exchanged. A first commenter pointed out what he saw as a flaw in the reasoning behind the original bill: the idea that people have a more basic right to a road than to a good living environment and protection of their property. Because the phrasing of the comment was not terribly clear, another commentator asked the community: "I wonder what the comment wants to communicate?" A third, seemingly rather well-informed commenter replied:

Probably that if one wants to have a snowmobile route, so the property rights and legal rights of thousand landowners can be violated in the right-based consideration. That doesn’t fulfill the first article in the European Human Rights Agreements’ first supplemental document.

In the discussion that followed, another participant introduced the large number of snowmobiles as an argument for rights-based consideration, that is, for putting the rights of snowmobilers to have a new road above the rights of other people. Then the original commentator responded:

Legal gimmicking is not the objective. Routes and their usage affect a large group of people, and they have environmental impact too. Right-based consideration is not applied in building roads either. There is a large group of stakeholders involved in route projects, the local conditions vary, the routes have to start and end somewhere, etc. And the snowmobile routes are intended to be built almost like roads.

Here the exchange of arguments is complicated by the fact that the initial commentator is not as articulate or clear as might be desirable. The community nonetheless makes an effort to rephrase the argument contained in the original post and continue the conversation on the basis of what they saw as the "probably" correct interpretation of the comment. The deliberation is about a complex issue: that of arbitrating between the rights of the snowmobilers’ community to free circulation and the property rights and rights to a safe environment of other stakeholders. The interaction brings in crucial pieces of information, such as the fact that the expired bill, which privileged snowmobilers’ rights, might have been in conflict with provisions in the European Human Rights Agreement or the fact, as pointed out by the initial commentator in his reply that the rights-based approach is not used in the decision to build actual roads (for cars). This fact allows him to make the a fortiori argument that there is no reason why the rights-based approach should then be used in determining where to build routes for off-road traffic. Beyond the clear statement of evaluative
preferences, this particular thread illustrates a discussion that aims to produce “reasonable, well-informed opinions” (Chambers, 2003, p. 309). Though no one seems to actually change their mind, the conversation reads as one in which participants are at least “willing to revise preferences in light of discussion, new information, and claims made by fellow participants” (Chambers, 2003, p. 309).

The tone of the conversation is a bit less respectful than in the previous exchange, as the original commentator sounds irritated and dismisses the legal distinctions introduced by others as “legal gimmicking.” Nonetheless the tone is civil enough to keep the conversation going and the other participants seem to make a genuine effort to listen to each other.

An interesting characteristic of online deliberation that transpires in those examples and others is that, in discussion threads, the deliberators often come and go such that the arguments exchanged are not necessarily between the same persons. Deliberation takes place between viewpoints rather than persons themselves. Anyone can show up on the thread and take up a view or attack it. This partial depersonalization of the exchange arguably allows it to be more fluid and enduring than actual face-to-face exchanges involving physical persons. The conversation can continue over several days, weeks, or months, with comments arriving at various times. In the crowdsourcing process we studied, deliberation happened both between the same individuals throughout the process and between new-comers and single-time visitors.

*Freedom, Equality, and Publicity*

Now, did the deliberation observed qualify as “democratic” in the broad sense that it took place between “free and equal” individuals and was of a “public” nature? We argue that it does.

The “freedom” component of this deliberation can arguably be taken for granted in a country like Finland and in a process that no one was forced to join. One may perhaps want to nuance this conception of freedom to reflect something like Rawls’ “worth of liberty” ([1971] 1999, p. 179) or the equal capacity to take advantage of the opportunity offered by the government. It is possible that by that standard the Finnish process wasn’t perfectly “free.” Still, of all countries in the world today, by almost all existing standards, Finland is one of the freest as well as one of the most educated and digitally connected (90 percent of the population has access to the Internet). If self-selected participation in an online deliberation crowdsourced by the government could not count as “free” in a minimal way there, it wouldn’t count as free anywhere.

A more positive form of freedom was further ensured on the platform by the publicity of the exchanges or what might be called the “horizontal transparency” of the deliberation (everyone’s comments were equally visible to all). Publicity was thus assured through the clear expectation and awareness that the exchanges were visible to the rest of the community. This public nature of the interactions between participants and the moderation of the interactions in turn ensured a minimal amount of civility and respect between the participants. As a
result, everyone was in theory free to contribute or speak their minds as they saw fit.

Equality of the participants is a complex notion. Equality in the sense of background political equality is something that, like freedom, can be assumed at least as a starting point in a democratic country like Finland. Additionally, formal equality of access to the process can also be assumed in a minimal sense. A thicker sense of equality would require equality of resources to influence the process, such as equality of education or self-confidence. Such an interpretation of equality as equality of influence or power would, however, go beyond what deliberative democrats have in mind when they talk about exchange of arguments among “equals.” We follow Knight and Johnson (1997) in embracing the simple standard of “equality of opportunity for influence.”

We believe that the observed exchanges met the standard of equality in that, first, we did not observe any leader or authority figure emerging over time. Second, the participants themselves perceived their experience of and interactions on the platform as broadly democratic and egalitarian in nature, where their definition of democracy ranged from “having power” to “having equal opportunities to make a difference” to “having a say” to “majority rule” and most included an egalitarian component. The same definitions came up when participants described their experience of the crowdsourcing process.

Additionally, as Figure 3 shows, 60 percent of the surveyed participants perceived that contributors spoke with an equal degree of authority, that is, were seen as having an equal right to speak. The reason why 40 percent saw this differently might be due to a possible ambiguity in the interpretation of the concept of authority. Some may have understood it as stemming from a right to speak, whereas others may have interpreted it as stemming from a competence or confidence to speak. Finally, to the extent that equality translates into respect toward others, evidence for perceived equality (by the participants) is the degree of mutual respect and civility that was both observed and experienced on the online platform.

On that latter point, two things are worth noting. First, the deliberation we observed on the platform exceeded the definitions of democratic deliberation we started from and met in part the additional criteria of mutual respect and civility (as per Gutmann & Thompson, 1996) and even sincerity (e.g., Kies, 2010). A small majority (51 percent) perceived other participants as sincere in their opinions and arguments, indicating a minimal but real amount of trust among the members of the community, of which most participated anonymously, using nicknames. In terms of respect and civility, we note that only about 20 comments were removed in postmoderation because of their inappropriateness in tone.

Second, it is worth emphasizing that participants perceived civility and respect in their interactions in spite of the fact that the platform allowed for anonymous comments and the people engaging actively tended to have strong opinions and intense preferences—a combination that often results in unpleasant exchanges. The strength of opinions was visible in the online interactions, in the interviews, and in the survey. A majority (64 percent) of the participants said that
they held strong opinions, as shown in Figure 3. The level of civility was surprising to the participants themselves, as illustrated in the following excerpt from an interview with a participant in the end of the crowdsourcing process (1):

And another thing that surprised me is that [. . . ] it’s been so level-headed there that either the moderator has been working hard or the people who are discussing are such that the worst excesses and gaffes are avoided. The discussion is now on a rather sober level instead of how internet discussions often are, this kind of horrible off-the-cuff remarks. (1, online participant)

This is not to say, however, that the overall level of respect and civility was high. In fact only a tiny majority (51 percent, of the survey respondents) perceived the tone of the exchanges on the platform as generally civil and respectful. By contrast, one-third of the respondents perceived the exchanges as lacking in civility and respect. These polarized perceptions may reflect the range of tones in the threads the users participated in. If a user joined a thread with a negative tone, they experienced a less civil interaction. If the tone was more positive, so was the experience.

One indicator of democratic deliberation along which the observed exchanges turned out to be the weakest in our experience, however, is “critical listening.” As shown in Figure 3, the majority of the survey respondents (56 percent) perceived that the participants did not make a strong effort to hear and understand others’ viewpoints. The reason for this could be the nature of online interactions, where participants tend to rush to express their views rather than take the time to show that they read others’ comments and take those into account. This weakness could arguably be resolved by design, at least partially. If the crowdsourced process were designed to foster listening and taking others’ comments into account, as opposed to just gathering knowledge, participants’ experience would likely be different. Another reason for participants’ disappointment with critical listening may have to do with what might have been unrealistically high expectations about other people’s ability and willingness to listen in that particular context.

All in all, even though the democratic deliberation was not of the highest quality, at least in terms of meeting all or most demanding standards of some deliberative democrats, it still qualifies as democratic deliberation according to the well-established definitions we committed to earlier. At any rate the exchanges we observed certainly made an epistemic contribution (getting all the useful ideas out into the public and refining them deliberatively). They could arguably additionally make a legitimating contribution (if the laws eventually reflected some of the considerations evoked in the crowdsourcing deliberation, which may still happen).

We propose, finally, to use the term “crowdsourced deliberation” to characterize further the democratic deliberation that emerged in the Finnish experiment. As we see it, crowdsourced deliberation is a specific variety of online deliberation (as defined by, e.g., Coleman & Shane, 2012, p. 3) or what others call “online public consultation” (e.g., Davies & Gangadharan, 2009). To the extent
that it meets the criteria of democratic deliberation, crowdsourced deliberation can be democratic, as it was in our experiment. But it need not be, which is why we offer crowdsourced deliberation here as an independent concept, possibly of use in different contexts, including nondemocratic ones.

More specifically, we thus define crowdsourced deliberation as a unique type of deliberation, in which the participants are self-selected and the crowdsourcing effort is embedded in a larger process whose outcome will be determined by authorities distinct from the crowd itself (e.g., government officials). The self-selected nature of the participants distinguishes crowdsourced deliberation from other types of online deliberation, in which participants are either randomly selected (e.g., Fishkin’s [2009] Deliberative Polls) or selectively chosen by organizers. The fact that the crowd is not the ultimate decision maker further distinguishes crowdsourced deliberation from deliberation whose outcome is meant to be binding. As with crowdsourcing more generally, power indeed remains with the crowdsourcer (here government officials), who controls how and when crowdsourcing happens, and how the crowd-generated input is used (Aitamurto, 2016). Finally, because of its distributed, asynchronous, and depersonalized nature, crowdsourced deliberation is more adequately described as taking place at the “deliberation system” level (Parkinson & Mansbridge, 2013) and qualifies as a form of “networked communication” among multiple actors in the political system (Coleman & Shane, 2012, p. 13).

Objection From (Lack of) Representativeness

A common objection usually raised against processes relying exclusively on self-selection, is to what extent was the group of participants representative of the larger population? To what extent was the deliberation truly inclusive of diverse views and thus “democratic” in the sense that goes beyond formal equality of participation?

Based on the demographic data in the survey, our group of participants has no claim to represent the larger public in any meaningful statistical sense. Among other unrepresentative characteristics, the participants were mostly male (over 80 percent of the survey respondents), mostly educated, and politically somewhat active. About one-third of the participants had written to a Member of Parliament, and one-third had written op-eds for newspapers. However, the majority (about 70 percent) had not, so the participants were a mix of those who already were civically active, and those who were less active. The majority of the participants (72 percent) had expressed their views on online forums like newspapers’ commenting sections at least once in the past five years, thus demonstrating familiarity with online participation. The participant population was thus skewed toward those who were already familiar with online participation. So the equality among the members of the participant crowd in fact conceals a rather profound inequality between them as a group and the rest of the country, who did not participate.

In response to this valid objection, one may want to emphasize that no democratic innovation is ever perfectly descriptively representative because none
of them are without some element of self-selection. In most classical deliberative settings outside of mini-publics (e.g., Town Hall meetings, participatory budgeting), participation is purely based on self-selection. Although all are included, in the sense of having an opportunity to join if they so desire (it is up to them in the end), not all end up having a say. Similarly here, although all are invited, only a few chose to join. Perhaps the element of self-selection is more important in our experiment than most other comparable deliberative settings but there are some mitigating elements to consider.

Second, one could argue that the important principle from a democratic standpoint is that everyone has an equal opportunity to say something and be heard. One would need to demonstrate of course that equality of opportunity is real and substantive, as opposed to just formal. From that point of view crowdsourcing experiments are bound to reflect the power and opportunity inequalities of existing societies and cannot by themselves be expected to remedy them. The point remains that to the extent that there is real equality of opportunity for participation in crowdsourcing experiments, the existence of a skewed sample of participants should not necessarily be seen as a problem. An interesting finding in our experiment is that the more active participants were at least sometimes seen as representing the views of the more passive ones. One reason for passive rather than active involvement once people had visited the site at least once was indeed the feeling that other people had already voiced their concerns, as illustrated in the following excerpt from an interview with a participant (14):

I can’t think of any set of issues that would have been left undiscovered. The discussion progressed rather well without me, from what I’ve seen.

(14, online participant, male)

Similarly, when prompted to comment on the 10 percent active participation rate on the website (about 700 registered users out of 7,000 visitors), an interviewee (1) pointed out that from his point of view, the opinions on the website represented “a rather good sample of the opinions that are in the air, even though there are only 700 active participants.” He further speculated that if the other 6,300 others “felt that their opinions had been greatly insulted, they probably would’ve become active too.” It is possible that when people feel that their concerns are represented or voiced, they self-consciously opt out. Observed passivity should thus not necessarily be interpreted as indifference, laziness, incompetence, or disapproval but possibly as a form of tacit consent to what other people are doing, at least when equality of access to the process is real. This would tend to suggest that although only a few people typically participate actively, their activities may be implicitly authorized and thus granted some degree of representative legitimacy by the other more passive participants.

Third, and more essential to legitimacy than statistical representativeness per se is issue representativeness, that is, the extent to which the deliberations were representative of the diversity of opinions on the topic existing in the larger community of affected interests. The interview data indicate that the
Crowdsourced deliberation reflected a more diverse array of opinions than the statistical composition of our group would suggest, thus showing representation of relevant issues. Interviewed participants reported being impressed by the inclusiveness and representativeness of the process in that respect, even if it was remarked that some groups were not represented (e.g., the indigenous Sami people, and people without access to the Internet, a small but important minority in Finland).

Fourth, inequality of representation would be problematic if it translated into an inequality of influence over the final policy outcome. In crowdsourcing, however, this problem is always considerably mitigated by the mediation of representatives, the ultimate decision makers, and that of researchers and bureaucrats, whose analysis of the data may in part correct for the skew. Because crowdsourcing is not a mode of direct decision making by the participants, the inequality is problematic only to the extent that it affects the online exchanges and may induce in participants the feeling that the deliberation is not truly representative or privileges one group over others.

Finally, even if the question of statistical representativeness were central to the legitimacy of processes, it is arguably a norm that should apply at the systemic level, not the local one; or to institutions aiming to represent the whole country (such as the legislature), not those aiming, more modestly, at gathering and injecting new ideas into a legislative process gone notoriously stale.

Conclusion

The questions that guided this research were: Are crowdsourcing platforms and processes conducive to democratic deliberation? How does the self-selected nature of the participation affect the democratic legitimacy of such processes?

In our assessment, crowdsourcing in the Finnish experiment was conducive to some degree of democratic deliberation, even though, strikingly, the process was not designed for it. Crowdsourcing in policymaking, thus, even when intended for knowledge and idea search only, is able to foster democratic deliberation. We proposed to use the term “crowdsourced deliberation” to describe the kind of deliberation occurring in the context of a crowdsourcing process, whether it is intentionally obtained or merely happened (as in our case). The deliberation we observed in the Finnish experiment thus combined both “democratic” and “crowdsourced” properties and can be characterized as “crowdsourced democratic deliberation.” We believe that if crowdsourced democratic deliberation were properly incentivized and rewarded, its quantity and quality would most likely increase.

An additional question one may want to ask is whether crowdsourced deliberation has the potential to turn into “mass deliberation” of the kind that mini-publics have given up on (see Fishkin, 2009). Crowdsourced deliberation has, in our view, such a potential, both because of its open nature (anyone can self-select to participate) and because of its asynchronous, depersonalized, distributed nature. Compared to mini-publics, the number of people that could be involved in crowdsourced deliberation, is, in theory, limitless. A smart design
of crowdsourcing technologies and the processes can facilitate large numbers of participants in meaningful ways.

The notion of lack of descriptive representativeness—a “feature” of crowdsourcing in general, rather than a bug—is endemic to participatory experiments that rely solely on self-selection, whether the experiment takes place on- or offline. Greater inclusiveness and absolute levels of participation need to be fostered in crowdsourced policymaking, to reach maximum inclusiveness in the process.

However, the democratic deliberation likely to occur in crowdsourcing is most certainly bound to remain skewed in a way that deliberative democrats should reconcile themselves with. After all, compared to the traditional lawmaking process, which only relies on civil servants in the Ministry and the stakeholder groups (interest groups in particular), a crowdsourced lawmaking process is more inclusive in the sense that ordinary citizens, as opposed to the usual suspects, are given a chance to become part of the knowledge-base creation for the law reform. Even if, in this particular Finnish experiment and, it seems, at this stage of the experimentation with crowdsourcing more generally, it is often well-educated white males who participate the most and in the largest numbers, this group is still injecting fresh ideas into the system. Furthermore, there is no telling what future and differently designed crowdsourcing processes on different laws (e.g., around breast-feeding in public rather than snow-mobile regulation) would generate in terms of women and minorities’ participation. Finally, the method of analysis of the crowd’s input can ensure that the self-selection induced skew in expressed preferences and arguments does not directly impact the policy output, which ultimately remains in the hands of parliament and elected representatives.

This case study has limitations. The study is also limited to one country and one issue. To address these limitations and test the generalizability of the results, future research should replicate crowdsourcing experiments in several countries and for other issues. Research questions worth exploring in the future are the quality and quantity of deliberation fostered in crowdsourcing and whether it is accompanied by peer-learning, preference-transformation, and consensus building. Finally, we suggest that scaling up crowdsourced policymaking to hundreds of thousands of people through the use of innovative technologies and natural language processing systems is a distinct possibility and most likely a part of the future of crowdsourcing.

**Tanja Aitamurto, Ph.D.,** Postdoctoral Fellow and Deputy Director, Brown Institute for Media Innovation, School of Engineering, Stanford University, Stanford, CA [tanjaa@stanford.edu].

**Hélène Landemore, Ph.D.,** Associate Professor of Political Science, Yale University, New Haven, CT.

**Notes**

1. See also Mercier and Landemore (2012), Landemore and Mercier (2012) for the emphasis on reasoning.
2. See Chambers (2014) for a defense of the compatibility of rational (or “neo-Kantian”) deliberation with rhetoric and emotions, against the interpretation of so-called sentimentalist (e.g., Garsten, 2006; Krause, 2008; Frazer, 2010 or Morrel, 2010) and naturalist (e.g., Connoly, 2002; Thiele, 2006) critics.

3. These additional criteria are not as essential in our view. Consensus, at any rate, is no longer seen as the normative horizon of democratic deliberation by most deliberative democrats (see Landemore & Page, 2015).

4. Mutz (2006) for example, equates deliberation with mere “cross-cutting exposure”—that is, exposure to conflicting viewpoints through political talk. For a critique of cross-cutting exposure as falling below the threshold of democratic deliberation see Landemore (2014).

5. We exclude from our definition one of Gastil’s (2000) components, namely “earnest decision making” as the production of an actual decision need not in our view be made part of the definition of deliberation. The decision might well ultimately be taken by means of a vote without affecting the value and completeness of the preceding deliberation.

6. www.maastoliikennelaki.fi

7. A new government was elected in Finland in Summer 2015, and it is unknown if and how the new Ministers will continue the law reform. As of early 2016 no progress can be reported.

8. The rate is in line with most Internet surveys. A meta-analysis of 39 studies found the un-weighted average response rate to be 34 percent for Web surveys (Shih & Fan, 2008).

9. In the survey, “authority” refers to equal authority to speak. Unlike in English connotations, it does not refer to formal authority, nor to the potential of reasoned elaboration. In this context, authority refers to an equal possibility of having a say.

10. We credit much of the civility to the official nature of the process—which was initiated by institutional authorities: A Ministry and the Parliament—and its horizontal transparency, which clearly prevented much of the nastiness often observed on Internet platforms that allow for anonymous contributions.

11. It would be interesting to assess the quality of the deliberative exchanges we observed using a different metric (e.g., that used by Karpowitz & Raphael (2014) to assess the deliberative quality of the average public hearing in the United States or that used by Steiner, Bächtiger, Spörndli, & Steenbergen (2004) to assess the quality of debates in the British House of Commons). We suspect the exchanges we observed would come out as more deliberative than at least the average public hearing in the United States.

12. We thank an anonymous reviewer for suggesting all these excellent points.

13. Even Fishkin’s (2009) Deliberative Polls, whose design is the closest to producing a perfect mirror-image of the population, tend to underrepresent busy wealthy people for whom the offered financial compensation for a weekend of deliberation is not enough of an incentive.

14. One may of course further wonder whether the people who do not even log in once on the platform are represented in any way, a question we are not able to answer at the present time.

References


