Self-selection In Crowdsourced Democracy: A Bug Or A Feature?

Tanja Aitamurto
School of Engineering,
Stanford University
475 Via Ortega, Stanford,
California, 94305
tanjaa@stanford.edu

Jorge Saldivar Galli
Department of Information
Engineering and Computer
Science, University of Trento
Via Sommarive 9, Povo, 38123
jorge.saldivargalli@disi.unitn.it

Juho Salminen
Lahti School of Innovation
Lappeenranta University of
Technology
Saimaankatu 11, LaHti, 15140
juho.salminen@lut.fi

ABSTRACT
This paper examines the tension between the norm of equal representation in democracy and the ideals of crowdsourcing, which strive for a large undefined, self-selected crowd. We argue that crowdsourcing, when used in democratic processes, can never meet the standard for statistical representativeness, which is the often-strived standard in democratic processes. We also argue that crowdsourcing shouldn't strive for statistical representativeness of the public, otherwise the virtues of crowdsourcing would be compromised and its benefits in crowdwork would not be achieved.

Author Keywords
Crowdsourcing; democratic innovations; representativeness

INTRODUCTION
Crowdsourcing in democratic processes, such as in public policy-making, has become more common in the recent years. Crowdsourcing has been used in knowledge search for law-reforms [1, 2], in policy-making in local governance [3] and in eliciting ideas for state policy agenda [4]. Crowdsourcing for democracy is a democratic innovation [5] in that it engages citizens in policy-making and brings them closer to the political decision-making power. An often-heard objection against crowdsourcing in democratic processes is the lack of representativeness of the participant crowd, which leads to input from a biased sample of population. This can be problematic in the context of democratic processes, because the crowdsourcing initiative can attract only certain demographic groups, or groups that share a certain political view. Therefore, it is justified to pose the following question: How does crowdsourcing as a participatory mechanism fit to the democratic ideal of equal representation? In this paper we address that question.

CSCW '14, February 15–19, 2014, Baltimore, Maryland, USA.

EQUAL REPRESENTATION, CROWDSOURCING AND SELF-SELECTION
For the purposes of this paper we consider equal representation consisting of two complementary aspects: each potential participant should have equal probability of participating, and each participant should contribute an equal amount. In polls and surveys, which are aimed to support democratic decision-making, representativeness is ensured by using representative sampling. A representative sample is a selected subset from a population that reflects the relevant features of the population accurately [6]. For example, if the population consists of 50% females, also the representative sample from that population should have about 50% females [7]. A form of random selection procedure is usually employed in sampling.

To ensure equal and comparative contributions, surveys for instance, aim to collect exactly the same amount of information from each participant [8]. Combination of random sampling and fixed contributions allows generalization of findings from the sample to the population as a whole. The deliberative poll [16] is a democratic innovation, which employs the idea of equal representation. It relies on the notion of "mini-publics", i.e. random samples of citizens, who are invited to deliberate about societal issues. Their opinion is measured after the deliberation and is considered to be a representative opinion about the issue, because it is assumed to represent the views of the larger public.

This traditional survey approach contrasts with crowdsourcing, in which participants can contribute as much or as little as they wish. Therefore, the participation in crowdsourcing is ruled by highly unequal levels of contributions, when most of the participants contribute very little and a few active ones contribute a lot [8]. This phenomenon of unequal participation has been first featured by Horowitz, who found that the participation of the crowd in content-generation online sites, such as YouTube, Wikipedia and Yahoo Groups, is governed by the rule of 1%. The rule of 1% means that out of every 100 visitors only 1% of them will create new content, of the remainder, 10% will refine and improve existing content while 89% will just consume it [9].

Crowdsourcing, one form of online content-generation, is an open call for anyone to participate in a task online [10, 11, 12]. It has been used to engage people from urban planning [13] to new product design and solving complex scientific problems [14]. A crowdsourcing system invites a crowd of people to help solve a problem defined by the system owners [11]. The "crowd" here refers to an undefined collection of people who participate in the open call and the system owner or task initiator, i.e., the crowdsourcer, can be any given entity, whether it is a company, institution, nonprofit organization or an individual. The term crowdsourcing is also used in contexts in which the task is open only to a restricted group like employees in a certain organization. Technology companies, such as IBM and Microsoft, have been using crowdsourcing to harvest ideas from employees to fuel grassroot innovation processes within their companies [15].

When used in democratic processes, like policy-making, crowdsourcing turns into a method for democratic innovations [5]. Democratic innovations engage citizens in democratic processes between the elections. Democratic innovations involve a variety of methods ranging from deliberative polls [16] and citizen assemblies to crowdsourced policy-making. In crowdsourced policy-making the citizens are invited to share their knowledge and ideas for improving the policy. Iceland, for instance, used crowdsourcing in its constitution reform in 2011 [2], and Finland crowdsourced its off-road traffic law reform in [17]. Crowdsourcing has also been used to extend the governments' capacities in data processing and analytics. For instance, the State of Minnesota has run an initiative called Minnesota's Citizen Lake in which citizens were invited to analyze and monitor the quality of the state's water resources [18].

Crowdsourcing has also been used for searching solutions for complex societal challenges, such as predicting solar flares. In this type of innovation challenges the participants propose innovative solutions for solving the problem. The SAVE Award launched by the White House is an example of employing the knowledge and skills of an online community to address a public concern, which in this case is reducing the public budget [19].

In contrast to random sampling used in polls and surveys, in crowdsourcing the participants self-select to work on the solution to the problem defined by the crowdsourcing system owner [20]. This means that the participants are not invited randomly to participate, but they initiate the participation themselves, leading to a biased sample — the crowd is not a representative sample of the larger public. Crowdsourcing thus collides with the notion of equal representation.

This contrast is illustrated in Figure 1, which presents the rank-ordered distributions of the amounts of information contributed in a survey using random sampling and in a crowdsourcing system. The diagram on left illustrates the distribution of contributions from a survey using random sampling and fixed contributions [21]. The figure on right illustrates the distribution of participant contributions to a crowdsourcing system [17].

In order for crowdsourcing to work the crowd needs to be large and at least some participants need to be knowledgeable and motivated to self-select to participate in problem solving [20]. At least in crowdsourced scientific problems the best solutions tend to come from people in technical and social marginality, who, thus, have different perspectives and problem solving approaches compared to the majority of the participants [22].

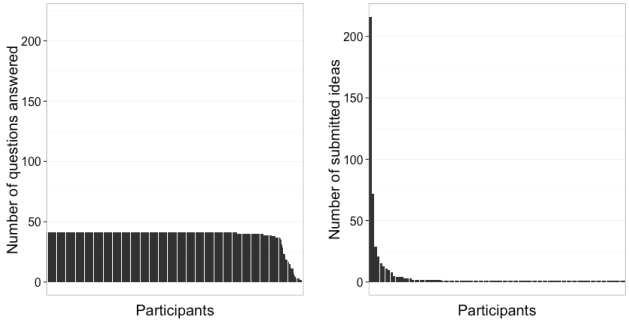

Figure 1. Comparing contributions between equal representation (fixed contributions in survey) and in crowdsourced knowledge search.

Altogether, crowdsourcing relies on the self-selection of participants, leading to highly unequal levels of contributions. Many of the best contributions come from the most unrepresentative individuals in the crowd. The contrasts between equal representation and crowdsourcing are outlined in Table 1.

	Equal Representation	**Crowdsourcing**
Participant Selection	Random	Self-selection
Contribution	Fixed	Unequal
Typical rank-order distribution of contributions	Uniform	Power-law
Goal	Generalization from sample population	Finding non-typical individuals that provide non-typical solutions
Outcome	"The public opinion" (poll result) /Majority result	Innovative solutions/Aggregated knowledge

Table 1: Comparison of equal representation and crowdsourcing

Because of the tension between the nature of crowdsourcing as a method based on self-selection and the strive to equal representation, an often heard objection against crowdsourcing in democratic processes is the lack of statistical representativeness of the participant crowd, which leads to biased samples. In democracy this can be problematic because the crowdsourced process may attract only certain demographic groups, or groups that share a certain political view. It is deceiving to consider the crowdsourced input as the public opinion.

CONCLUSION

In this paper we have exposed the tension between the ideals of equal representation and self-selection in crowdsourcing. Crowdsourcing aims for large, diverse participation, which is based on self-selection, and those are the virtues of crowdsourcing. Equal representation, instead, presumes equal representation (in the form of statistical representativeness of the public) and equal contributions. If crowdsourcing aims for equal representation, its virtues are compromised and that can undermine the method's advantages. Therefore, crowdsourcing, even when applied in democratic processes, shouldn't attempt to follow the norm of equal representation. Instead, crowdsourcing as a method for participatory democracy should be cherished by enhancing its virtues and developing the method using those virtues, whether crowdsourcing is used in distributed group-work or for large online crowds.

REFERENCES

1. Aitamurto, T. Crowdsourcing for Democracy: New Era In Policy Making. *Publications of the Committee for the Future, Parliament of Finland.* 2012.
2. Landmore, H. Inclusive Constitution–Making: The Icelandic Experiment. *Journal of Political Philosophy*. 2014
3. Brabham, D. C. Crowdsourcing the public participation process for planning projects. *Planning Theory,* 8, 242–262. 2009
4. Nelimarkka, M., Nonnecke B., Krishnan S., Aitamurto T., Catterson D., Crittenden C., Garland C., Gregory C., Huang C.C., Newsom G., Patel J., Scott J., Goldberg K. Comparing Online Civic Engagement Interfaces using the "Spectrum of Public Participation" Framework. *Conference on Crowdsourcing for Politics and Policy*. Oxford. 2014
5. Smith, G. Democratic Innovations: Designing Institutions for Citizen Participation**.** *Cambridge University Press.* 2009
6. Bryman, A. Social Research Methods. *Oxford University Press.* 2008.
7. Cherry, K. What is a representative sample? *Psychology.about.com.* Retrieved from: *http://psychology.about.com/od/rindex/fl/What-Is-a-Representative-Sample.htm*
8. Salganik, M., and Levy, K. Wiki surveys: Open and quantifiable social data collection. *arXiv preprint arXiv:1202.0500*. 2012
9. Horowitz, B. Creators, Synthesizers, and Consumers. 2006. *Retrieved from: http://blog.elatable.com/2006/02/creators-synthesizers-and-consumers.html*
10. Brabham, D. C. Crowdsourcing as a model for problem solving an introduction and cases. *Convergence: the international journal of research into new media technologies*, 14(1), 75-90. 2008
11. Doan, A., Ramakrishnan, R., Halevy, & A. Y. Crowdsourcing systems on the world-wide web. *Communications of the ACM, 54(4)*, 86-96. 2011
12. Howe, J. Crowdsourcing: How the power of the crowd is driving the future of business. *Random House*, 223-246. 2008.
13. Brabham, D. C., Sanchez, T. W., & Bartholomew, K. Crowdsourcing public participation in transit planning: preliminary results from the next stop design case. *In TRB 89th Annual Meeting Compendium.* 2010
14. Aitamurto, T., Leiponen, A., & Tee, R. The Promise of idea crowdsourcing—Benefits, contexts, limitations. White paper for Nokia Ideas Project. 2011.
15. Bailey, B., and Horvitz, E. What's your idea? a case study of a grassroots innovation pipeline within a large software company. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010.
16. Fishkin, J. When the people speak: Deliberative democracy and public consultation. *Oxford University Press*. 2009
17. Aitamurto, T., and Landemore, & H. Democratic participation and deliberation in crowdsourced legislative processes: The case of the law on off-road traffic in Finland. *6th Conference on Communities and Technologies.* 2013
18. Nambisan S., and Nambisan P. Engaging Citizens in Co-Creation in Public Services: Lessons Learned and Best Practices. IBM Center for The Business of Government. 2013.
19. Long, E. Administration announces finalists in cost-cutting contest. GovernmentExecutive.com. 2009. *Retrieved from:* http://www.govexec.com/oversight/2009/12/admin

istration-announces-finalists-in-cost-cutting-contest/30478/

20. Afuah, A. N., and Tucci, C. L. Crowdsourcing as a solution to distant search. *Academy of Management Review 37(3)*. 355–375. 2012

21. Saunila, M. Innovation capability for SME success: *Perspectives of financial and operational performance, Journal of Advances in Management Research, Vol. 11, No. 2*, pp. 163-175. 2014

22. Jeppesen, L. B., and Lakhani, K. R. Marginality and problem-solving effectiveness in broadcast search. *Organization Science*: *Articles in Advance*, 21(5), 1016–1033. 2010