

Methodological and ethical implications of testing alternative designs for technologies supporting democratic processes

Mauro Cherubini

mauro.cherubini@acm.org

In recent years we have witnessed an increased number of solutions to support democratic discourse and enable citizen to exercise unmediated political power. For example: Liquid Feedback (or lqfb) is a software that was released in 2009 that implements a Delegated voting system which is to establish a new form of political representation and participation that takes into account the knowledge disparity of its participants¹. Similarly the Movement 5 Star, an Italian movement for direct democracy, recently released the M5S OS, a web portal that allows members to vote on law proposal currently under discussion in the parliament and to comment and decide on initiatives that might lead to new law proposals².

These softwares have the same core objective: in essence they aim at enabling users to express their political views. However, their interaction design is dramatically different and might lead to extremely different outcomes. For instance, while lqfb uses the *Schulze Method* to select the winner in case of preferential voting, other system like the M5S OS might use a ranked pair voting method.

1 See also <http://en.wikipedia.org/wiki/LiquidFeedback>

2 See also http://en.wikipedia.org/wiki/Five_Star_Movement

While from a sociological perspective both approaches could be analyzed with the same lens in order to understand why society is moving away from representative democracy, from a human-computer interaction perspective it is extremely hard to understand the different outcomes that such systems might yield.

These two softwares are extremely different. One of the most striking difference concerns how pools are organized. While in M5S OS one voter counts one vote only and the system does not allow delegation, lqfb has a proxy voting feature which allow the user to delegate his voting power to another user that s/he trusts. Delegation can be revoked at any time, it can be assigned to another user for a given topic or for a given ballot only. Hypothetically speaking, if we were to observe the exact same pool being conducted with the exact same voters over both systems, we would probably obtain radically different results. Then we would be able to reflect on the ability of each of these systems to allow participation in the democratic process and the ability of each outcome to represent the real political views of the citizens. Conducting such comparisons would be extremely important for the evolutions of these systems because we should aim for a design that *does not bias results* in any manner, that *represents with fidelity the political views* of the voters, that *allows the voters to take informed decisions*, and finally that *encourages participation*.

Unfortunately, conducting experimental comparisons is challenged by several methodological and ethical issues:

- *It is extremely complicated to control all variables in a live experiment.* The scientific method requires manipulation of one experimental factor while maintaining all other factors equal. Unfortunately, achieving this in a live experiment is extremely complicated if not impossible. For example, while one user might access the experiment with a 21" display thus allowing multiple sets of information, another user might use a 10" tablet which could potentially change the experience s/he might have on the voting platforms. Some of these variables might even be impossible to control because they might be related to the historical context at the time the test is run (e.g., the public opinion at the time of the experiment, etc).

- *A within-subject design can lead to presentation and recall biases, while a between subject design can lead to sampling biases.* If we ask users of a voting system to use two different variants of a voting system to cast their vote about a given topic, they might be influenced by the particular order in which they see the variants or their first experience voting might influence how they might vote while voting a second time on the same topic. Conversely, if we used two different groups of users to cast votes for the exact same topic using the two

variants of the design, we might inadvertently introduce a bias in the selection of participants assigned to the two variants.

- *There are ethical implications in applying experimental manipulations to real pools.* Similarly to the research conducted with/on embryonic stem cells, researchers working on technologies for democratic processes face the dilemma of influencing ballots that have real-life implications for the life of the citizens. For instance an experiment might be conducted when voters are debating how to amend the immigration law of the country. Depending on the design of the feature at hand voters can be nudged to vote in a given manner to the pool. In turn the biased results of the pool might result in immigrants being deported, or imprisoned.

Specific design choices might be accused by opposing parties to nudge users in a given direction and to bias the results, so how can we fully explore a design space while at the same time maintaining transparency with regard to the users and the general public? Conducting experimental comparisons of different design for technologies that support democratic processes is extremely important because it is the only way to demystify blunt criticisms, achieve transparency, and make progress on these technologies. It is our responsibility also, as designers and as researchers to make sure our solutions are effective in achieving a truly democratic interaction in our society.