The California Report Card: A Mobile Crowdsourcing Tool for Timely Public Policy Advice

Tanja Aitamurto[3], Daniel Catterson[3], Camille Crittenden[3], Chris Garland, Allen Ching-Chang Huang[1], Sanjay Krishnan[1], Matti Nelimarkka[2,4], Gavin Newsom, Jay Patel[1], John Scott[3], Ken Goldberg[1]

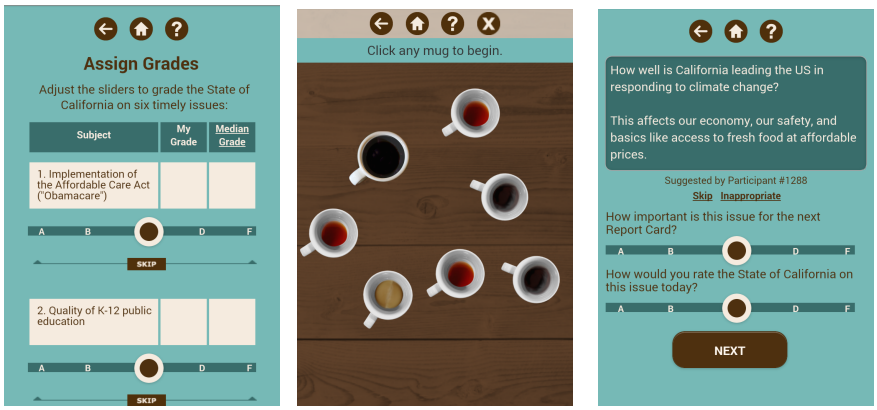
[1] University of California, Berkeley, Electrical Engineering and Computer Sciences [2] University of California, Berkeley, School of Information [3] The Center for Information Technology Research in the Interest of Society (CITRIS), University of California Berkeley [4] Helsinki Institute for Information Technology HIIT, Aalto University

Introduction

The California Report Card (CRC) is a mobile optimized web application that allows participants to advise the California state government on timely policy issues. This tool was developed in collaboration with the Office of the Lieutenant Governor of California.The CRC combines an online opinion poll, where participants grade the state on a number of issues (Fig. 1a), and a free-form textual ideation space where participants can suggest questions for future polls.

The ideation space is designed as a metaphorical café table, where participants discuss important state issues over a cup of coffee. We represent participant's suggestions as "mugs" on a coffee table, with your own mug placed in the center (see Fig. 1b). The mugs are placed using Principal Component Analysis (PCA) on the initial issues so mugs that are closer are ones who graded the state similarly. This approach has been previously evaluated in a controlled laboratory study [3].

The CRC applies the wisdom of the crowd to identify the most insightful suggestions. Participants evaluate each suggestion on two axes: "How important is this for the next report card?" and "How would you grade the state of California on this issue today?" (Fig. 1c). We rank the suggestions by the lower bound of the 95 % confidence interval of the mean importance grades. Our prior results suggest that this reputation model is particularly effective at ranking free-form textual suggestion [1], and similar models have been applied in other domains [8, 5].



(a) Grading the state (b) Other participants' suggestions (c) Grading one spesific idea

Figure 1: Examples of the user interface

As of March 10th 2014, a total of 6,782 participants visited the CRC and submitted 22,281 grades. We have also collected 333 textual suggestions. In our analysis of the collected data, we address the following research questions:

1. A detailed analysis of how participants interact with the system.
   (a) Statistically significant participation differences between mobile devices and desktops.
   (b) Statistically significant participation differences between urban and rural participants.
   (c) Statistically significant differences in data entry and skipping.
2. As this is not a randomized survey, we provide an analysis of some of the biases introduced.
   (a) We test for population/selection bias by running a randomized reference survey asking the same questions as our application.
   (b) We hypothesis test social influence biases, such as whether partcipants were more likely to change their grades upon seeing the population median.

Related Work

Introducing mobile devices to political participation has been well studied: Rheingold [6] explores this topic and Hermanns [4] suggests mobile phones are impactful, but understudied area in ICT for political participation-research. Furthermore, mobile support for public participation has been explored in city planning acties, mainly as mobile platforms are location and context-aware [2, 7].

This work explores the concept of mobile devices for policy further: we explore both issue-representativeness of our sample and the differences between mobile and desktop users in the system.Similarly, exploring the biases in social media has recently received much attention [5]. The authors conducted a randomized experiment with Reddit.com to identify the effects of "social influence", whether seeing a higher score on a comment prompts future users to leave higher grades.

Data and methods

We analyze the data entered by the participants, and the application was instrumented to collect all participant interactions (eg. button clicks, premature exits, etc.). We determine the location of the participant using the entered zip code, and we additionally cross-reference these zip codes with US census records to group them into regions and urban vs. rural. We determined the device (mobile vs. desktop) by using the user agent entry in our webserver logs. We test our findings with significance tests, such as F-test, t-test, and chi-squared tests of correlation.

Findings

Based on our preliminary data, here are our findings:

1. **Mobile devices and desktops**
   Between mobile devices and desktops there was not a significant difference suggestion length in characters ($\bar{x}_m = 111.3, \bar{x}_d = 112.5$), however the difference in variances was significant ($SD_m = 102.8, SD_d = 87.1, p < 0.01$). Also with issue grading per user difference was not high ($\bar{x}_m = 5.9, \bar{x}_d = 6.2$), however the difference in variances was significant ($SD_m = 13.1, SD_d = 15.2, p < 0.01$).

2. **Urban vs. rural**
   On average rural participants evaluated more issues ($\bar{x}_u = 6.8, \bar{x}_r = 9.9$), however this difference is yet not statistically significant ($p = 0.14$).

3. **Self-selection bias**
   CRC allows participants to skip questions and dialogs. We found evidence of self-selection in the system; that is the submission of data is correlated with the values given. Participants who chose to submit emails addresses after participation had statistically significantly higher grades for the state on two out of six issues ($p = 0.048$, $p = 0.014$).

4. **Biasing effect of median grades**
   We tracked participants who changed their grades, and observed moderate tendency for the participant to move their grade towards the revealed population median. For all issues, the null hypothesis was very unlikely ($p < 10^{-5}$). We also tested to see if this was an artifact of the mobile

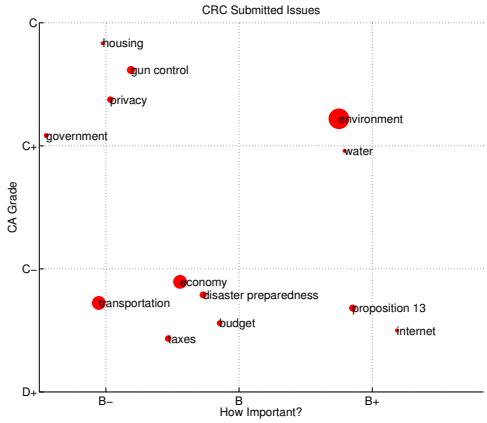


Figure 2: Topics raised by the participants

interface, and evaluated alternative models to see if they could explain the correlations.

5. **Reference survey** We have an in-progress randomized survey to evaluate how representative our sample is. We plan on applying a correction to align the means and variances of the two surveys to best explain towards which sub-populations we are biased.

Policy Outcomes and Future Work

Figure 2 shows popular topics suggested by participants and the ratings that they received. We work closely with the California state government to communicate these suggestions and trends. As this project is still underway, the analysis of the data and further policy outcomes are still expected to emerge. We plan on introducing a Version 2.0 later this year, which builds on the collected suggestions from Version 1.0.

References

[1] Ephrat Bitton. A spatial model for collaborative filtering of comments in an online discussion forum. In *Proceedings of the third ACM conference on Recommender systems*, pages 393–396. ACM, 2009.

[2] Morten Bohø j, Nikolaj G. Borchorst, Susanne Bø dker, Matthias Korn, and Pär-Ola Zander. Public deliberation in municipal planning. *Proceedings of the 5th International Conference on Communities and Technologies - C&T '11*, 28(2):74–82, June 2011.

[3] Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, pages 1175–1184, New York, New York, USA, 2010. ACM Press.

[4] Heike Hermanns. Mobile Democracy: Mobile Phones as Democratic Tools. *Politics*, 28(2):74–82, 2008.

[5] Lev Muchnik, Aral Sinan, and Sean J. Taylor. Social influence bias: A randomized experiment. *Science*, pages 647–651, 2013.

[6] Howard Rheingold. *Mobiilijoukot. Seuraava yhteiskunnallinen vallankumous*. Like-kustannus, Helsinki, 2003.

[7] Joanna Saad-Sulonen and Andrea Botero Cabrera. Setting up a public participation project using the urban mediator tool. In *Proceedings of the 5th Nordic conference on Human-computer interaction building bridges - NordiCHI '08*, page 539, New York, New York, USA, October 2008. ACM Press.

[8] Sean Wallis. Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods. *Journal of Quantitative Linguistics*, pages 178–208, 2013.