SurvRaid: Arbitrary Groupings for Increased Participation on the Web

Bryce Tham
Stanford University
Stanford, CA 94305, USA
bjtham@stanford.edu

Riley DeHaan
Stanford University
Stanford, CA 94305, USA
rdehaan@stanford.edu

Alejandrina Gonzalez
Stanford University
Stanford, CA 94305, USA
alegre@stanford.edu

Abstract
In this paper, we introduce SurvRaid, a survey technique that utilizes arbitrary groupings to address the problem of low participation rates in online settings in which participant affiliations may not be available a priori. Using the application domain of web surveys, we evaluated our technique by distributing a week-long lifestyle survey to just under 2000 students at Stanford University, dividing them up into arbitrary, affiliated, and control experimental sets. Our results suggest that arbitrary groupings can be as effective as affiliated groupings and outperform no groupings at all, though the results are not statistically significant. We end with recommendations for future work.

Author Keywords
Crowdsourcing; collaboration; survey methodology.

Introduction
Web-based surveys, while popular among researchers for their low cost and wide reach, are not without problems. These surveys often yield low response rates when compared to other data collection methods, reducing the quality of survey data due to non-response bias [2]. There has been significant research into improving response rates in web surveys, primarily regarding the use of incentives, but success has largely been mixed [7]. We propose a simple survey technique, SurvRaid, that utilizes arbitrary groupings

for web surveys to solve the problem of low participation rates. We supplement this grouping structure with a simple user interface to increase the salience of competitive standing between groups [4] [6]. The results of our study have implications for the design of similar systems requiring the “wisdom of crowds” and crowdsourcing more broadly in demonstrating a simple method for increasing engagement by arbitrarily grouping participants with minimal a priori knowledge of those participants. Potentially, it could even be used for larger-scale participation problems such as civic engagement.

Previous research in social psychology and prior literature on survey design suggest that a group structure might improve survey response rates. One author found mailed surveys perceived as gathering domestic data increased response rates by 3%-5% relative to surveys perceived as collecting cross-national data [5], demonstrating perceptions of meaningful groupings can affect response rates. Furthermore, according to the minimal group paradigm, people tend to show bias to benefit their own group, regardless of how arbitrarily those groups are formed [8]. Previous research on the minimal group paradigm has typically been applied to social contexts [3], with most of these studies involving individuals who interact with one another frequently. However, behaviors of in-group bias predicted by the minimal group paradigm have also been observed in more distant web-based contexts [1]. These findings suggest that the biasing behaviors of the minimal group paradigm could be formed even online, potentially to competitively motivate survey takers and increase survey response rates.

Our SurvRaid technique is designed to utilize minimal group paradigm effects to increase survey participation. In SurvRaid, participants are randomly placed in competing arbitrary groups for survey-taking, and group participation is communicated via reminders throughout the survey period to increase salience of grouping. A raffle is held among the respondents of the winning team. We describe our study evaluating the SurvRaid technique in this paper.

Study Design

We evaluated the SurvRaid technique in a between-subjects study in which over 2000 undergraduate and graduate students from the CS, EE, and Biology departments at Stanford University were randomly sampled as participants. We excluded ourselves from the study as well as others who had prior knowledge of our study. We also excluded students with majors in at least two of the departments we sampled since these students are not uniquely affiliated with one of the departmental groups.

The overall population of 1920 students was split between two week-long runs of our survey. We planned to include 960 students in each run, further divided into a control set of 240 students, three groups of 120 students affiliated by department (together referred to as the affiliate set), and three arbitrary groups of 120 students (together the arbitrary set). The groups in the affiliate set contained only students randomly sampled from the samples of the respective departments. The students from the groups in the arbitrary set contained students randomly sampled from the remainder of the 1920 students after sampling for the affiliate set. These group sizes were selected as near the largest group sizes supported by all of the departments after the exclusions (these departments were among the largest in the university). The survey contained general questions concerning student life and was designed to take around ten minutes to complete.

The arbitrary set was divided generically into a blue, orange, and red groups or "teams", and the affiliated set was

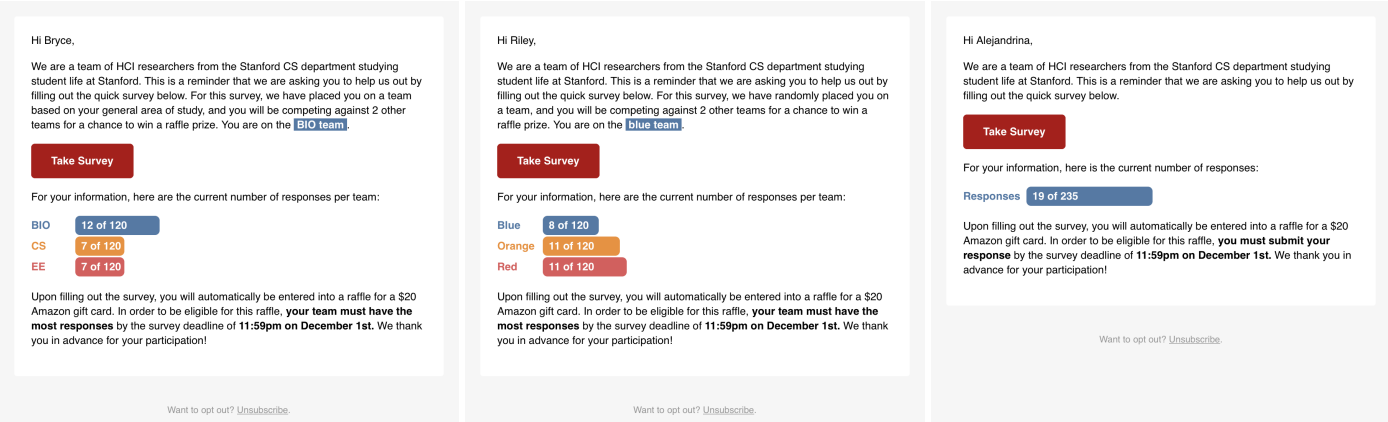

Figure 1: Sample reminder emails that were sent out to the arbitrary (left), affiliated (middle), and control (right) sets via Qualtrics.

divided into BIO, CS, and EE "teams". The survey request email content was designed to make clear the purposes of the teams, and an incentive of a $20 gift card was added to increase response rates. For the control set, the incentive was raffled to a randomly chosen survey respondent. For the arbitrary and affiliated sets, the incentive was raffled to a random respondent of the winning group. For the purposes of the raffle, the control set may be considered as a "single group" competition relative to the other sets with multiple competing groups. This incentive structure of a raffle containing the respondents of only the winning group may represent a potential confound for our results, but was included to make the groups more meaningful.

The surveys were administered with Qualtrics, which allowed for the emails to be personalized with participant first name. Typical survey reminder emails sent out to the different experimental sets are shown in Figure 1. The bar graphs of survey completion were included in the emails to increase the salience of the groupings. For each survey run, the initial emails were sent out at 8 a.m. on Day 1 and the reminder emails at 8 a.m. on Day 4 and Day 7, with the survey closing at 11:59 p.m. on Day 7. At the end of the study, a post-study survey was sent to respondents of the arbitrary and affiliated sets to learn about their perceptions of the team structure.

Results & Analysis

Our results were taken from two separate runs of our experiment, each spanning seven days. Of the 1920 total survey invitations sent out, 21 bounced (5 arbitrary, 7 affiliated, 9 control) and 5 (control only) did not have associated SUID email addresses to them, and thus these participants were excluded from our results. The final number of emails was 954 for the first run (360 arbitrary, 358 affiliated, and 236 control) and 940 for the second run (355 arbitrary, 355 affiliated, 230 control).

Response Rates by Set

In terms of response rates, our results indicate that the arbitrary set performed about the same as the affiliated set, and both the arbitrary and affiliated sets outperformed the control set, though the differences are not statistically significant as discussed below. In the first run, we received response rates of 17.78% (64/360) for the arbitrary set, 18.16% (66/358) for the affiliated set, and 12.29% (29/236) for the control set. The margin was much closer in the second run, with response rates of 15.49% (55/355) for the arbitrary set, 15.49% (55/355) for the affiliated set, and 14.78% (34/240) for the control set. See Tables 1 and 2.

Across the two runs, the overall response rates were 16.64% (119/715) for the arbitrary set, 16.83% (120/713) for the affiliated set, and 13.52% (63/466) for the control set. The differences between the affiliate set and the control set (χ^2 = 2.11, p = 0.15) and between the arbitrary set and the control set were not significant (χ^2 = 1.88, p = 0.17). The difference between the arbitrary set and the control set was also not significant (χ^2 = 0.0006, p = 0.98). The results are summarized in Figure 2. Throughout the study, both the arbitrary and affiliated sets saw higher response rates, cumulatively, across each day of the experiment. This is especially true for the first reminder day (Day 4) when the team rankings were first made salient, in which response rates increased by 4.9% and 4.76% for the arbitrary and affiliated sets respectively but only 2.8% for the control set when compared to the previous day. On the second reminder day (Day 7), the difference was much smaller, and the control set slightly outperformed the arbitrary and affiliated sets on the final day. One possible explanation for this may be that the group rankings have all but solidified with only one day remaining in the study, and so members of the losing teams in both the arbitrary and affiliated sets felt less inclined to submit a response. Indeed, for the arbitrary and affiliated

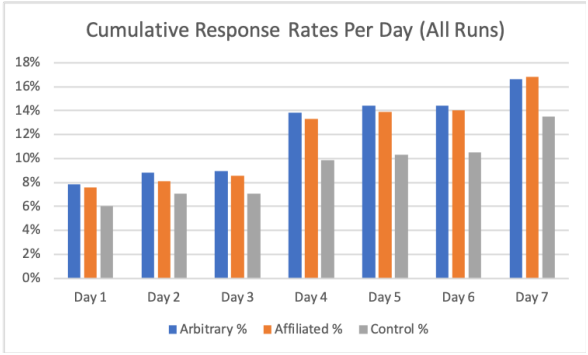

Figure 2: Overall cumulative response rates across both runs. The arbitrary and affiliated sets outperformed the control, but not statistically significantly so.

sets in both survey runs, the team with the most responses on the final day was the team currently in the lead. In the words of one respondent, “My ‘team’ was already leading so it looked like there was a fair chance we’d win and I’d have a shot at the incentive reward.”

Team Ranking as Motivation

In terms of team rankings, our results indicate that teams who performed well initially continued to perform well, and teams who performed poorly initially continued to perform poorly. In fact, for every arbitrary and affiliated set, the team with the highest number of responses after the first day also had the highest number of responses after the final day. In our post-survey results, nearly a third (about 30%) of respondents indicated that team ranking played some role in motivating them to complete the survey, suggesting that making this aspect more salient has some effect on increasing participation.

	Blue Team	*Orange Team*	*Red Team*	*Arbitrary Set*	*BIO Team*	*CS Team*	*EE Team*	*Affiliate Set*	*Control Set*
Day 1	11	10	11	32	13	7	11	31	11
Day 2	0	1	1	2	1	1	0	2	3
Day 3	0	0	0	0	1	1	0	2	0
Day 4	6	4	7	17	6	1	7	14	7
Day 5	3	1	0	4	3	0	0	3	1
Day 6	0	0	0	0	1	0	0	1	0
Day 7	4	4	1	9	6	3	3	12	7
Total Count	24	20	20	64	31	13	21	65	29
Response Rate	20.00%	16.67%	16.67%	17.78%	26.05%	10.92%	17.50%	18.16%	12.29%

Table 1: Number of responses per day, total number of responses, and response rates per team/set for the first run. Final number of emails sent: 360 arbitrary, 358 affiliated, and 236 control.

	Blue Team	*Orange Team*	*Red Team*	*Arbitrary Set*	*BIO Team*	*CS Team*	*EE Team*	*Affiliate Set*	*Control Set*
Day 1	7	8	9	24	11	6	6	23	17
Day 2	1	3	1	5	0	1	1	2	2
Day 3	0	0	1	1	1	0	0	1	0
Day 4	6	6	6	18	12	3	5	20	6
Day 5	0	0	0	0	1	0	0	1	1
Day 6	0	0	0	0	0	0	0	0	1
Day 7	1	1	5	7	5	1	2	8	7
Total Count	15	18	22	55	30	11	14	55	34
Response Rate	12.61%	15.38%	18.49%	15.49%	25.00%	9.32%	11.97%	15.49%	14.78%

Table 2: Number of responses per day, total number of responses, and response rates per team/set for the second run. Final number of emails sent: 355 arbitrary, 355 affiliated, and 230 control.

Limitations & Future Work

While the arbitrary and affiliated sets outperformed the control in our study, our results were not statistically significant. Further studies may require larger populations to better discriminate the effects of arbitrary groupings. A potential limitation of our study is that the affiliation that we chose (departmental affiliation) is relatively weak, as members of the same department are not likely to have known each other beforehand. In the post-survey, while some mentioned that being on a departmental team was meaningful, others disagreed. One participant stated, "I don't really feel like I belong to CS even though I'm working toward a CS degree. There's so many of us that it's hard to characterize us as a cohesive union of people." Had we chosen a stronger affiliation (like student dormitories or campus organizations), the results may have been different, though it should be noted that finding a sufficiently large group of participants with such strong affiliation is difficult. To this end, we propose a future work that might take into account affiliation strength and its effects on participation. Our study was also affected by the limited size of the departments, which total only a few hundred students. In real-life applications like crowdsourcing and public polling, the sample size could reach well over two thousand. It would be insightful to see how our results would scale to much larger sample sizes, and whether we would see statistical significance then.

Conclusion

Inspired by the minimum group paradigm, we devised SurvRaid, a technique that relies on arbitrary groupings to increase participation on the web. We evaluated SurvRaid by distributing a web survey to randomly sampled Stanford students and grouping participants into arbitrary, affiliated, and control sets, and our results show some promise that such groupings could lead to increased participation. Further work is needed to understand the impact that affiliation strength has on online participation, and our work has laid the foundation for future study.

REFERENCES

1. Yair Amichai-Hamburger. 2005. Internet Minimal Group Paradigm. *CyberPsychology & Behavior* 8, 2 (2005), 140–142. DOI: http://dx.doi.org/10.1089/cpb.2005.8.140 PMID: 15938653.
2. Colleen Cook, Fred Heath, and Russel L. Thompson. 2000. A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. *Educational and Psychological Measurement* 60, 6 (2000), 821–836. DOI: http://dx.doi.org/10.1177/00131640021970934
3. Michael Diehl. 1990. The Minimal Group Paradigm: Theoretical Explanations and Empirical Findings. *European Review of Social Psychology* 1, 1 (1990), 263–292. DOI: http://dx.doi.org/10.1080/14792779108401864
4. Robert M. Groves, Eleanor Singer, and Amy Corning. 2000. Leverage-Saliency Theory of Survey Participation: Description and an Illustration*. *Public Opinion Quarterly* 64, 3 (2000), 299–308. DOI: http://dx.doi.org/10.1086/317990
5. David Jobber and John Saunders. 1988. An Experimental Investigation into Cross-National Mail Survey Response Rates. *Journal of International Business Studies* 19, 3 (01 Sep 1988), 483–489. DOI: http://dx.doi.org/10.1057/palgrave.jibs.8490386
6. Oded Nov, Ofer Arazy, Kelly Lotts, and Thomas Naberhaus. 2013. *Motivation-Targeted Personalized UI Design: A Novel Approach to Enhancing Citizen Science Participation*. 287–297. DOI: http://dx.doi.org/10.1007/978-1-4471-5346-7_15

7. Eleanor Singer and Cong Ye. 2013. The Use and Effects of Incentives in Surveys. *The ANNALS of the American Academy of Political and Social Science* 645, 1 (2013), 112–141. DOI: http://dx.doi.org/10.1177/0002716212458082

8. Henri Tajfel. 1970. Experiments in Intergroup Discrimination. *Scientific American* 223, 5 (1970), 96–103. http://www.jstor.org/stable/24927662