Check for updates

Article

Field Methods
2016, Vol. 28(3) 231-246
© The Author(s) 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1525822X15603149
fmx.sagepub.com
SAGE

Comparisons of Online Recruitment Strategies for Convenience Samples: Craigslist, Google AdWords, Facebook, and Amazon Mechanical Turk

Christopher Antoun[1], Chan Zhang[2], Frederick G. Conrad[1], and Michael F. Schober[3]

Abstract

The rise of social media websites (e.g., Facebook) and online services such as Google AdWords and Amazon Mechanical Turk (MTurk) offers new opportunities for researchers to recruit study participants. Although researchers have started to use these emerging methods, little is known about how they perform in terms of cost efficiency and, more importantly, the types of people that they ultimately recruit. Here, we report findings about the performance of four online sources for recruiting iPhone users

[1] Institute for Social Research, University of Michigan, Ann Arbor, MI, USA
[2] School of Journalism, Fudan University, Shanghai, China
[3] The New School for Social Research, New York, NY, USA

Corresponding Author:
Christopher Antoun, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA.
Email: antoun@umich.edu

to participate in a web survey. The findings reveal very different performances between two types of strategies: those that "pull in" online users actively looking for paid work (MTurk workers and Craigslist users) and those that "push out" a recruiting ad to online users engaged in other, unrelated online activities (Google AdWords and Facebook). The pull-method recruits were more cost efficient and committed to the survey task, while the push-method recruits were more demographically diverse.

Selecting research participants for experiments based on their availability (convenience sampling) has served an important role in social science research for several decades as a low-cost alternative to probability sampling (Baker et al. 2013; Henry 1990). The place to find recruits for these samples, however, is rapidly changing. Traditionally, researchers found them by posting flyers in public places and placing print ads in newspapers and magazines. With the rise of online research, researchers are increasingly looking for recruits on the web.

In this article, we focus on four types of online recruitment methods: online classifieds, crowdsourcing, search engine advertising, and social media advertising. Each is described in Table 1. We divide these methods into two groups: those that "pull in" online users actively looking for paid work and those that "push out" a recruiting ad to online users engaged in other, unrelated online activities.

Pull applies to any strategy to target online users who are actively trying to complete tasks for cash incentives or rewards on the web. One pull method is online classifieds on sites like Craigslist to recruit volunteers for research projects. This was one of the first web-based methods used by researchers (e.g., Anderson et al. 2013; Worthen 2013). Recently, newly available online recruiting sources have vastly expanded the opportunities for researchers. In particular, crowdsourcing has become a widely used pull method. This allows researchers to find "workers" who are looking to complete surveys (among other tasks) for small amounts of money. One example is Amazon Mechanical Turk (MTurk), a crowdsourcing service site that has become increasingly popular for social science experiments (e.g., Berinsky et al. 2012; Crump et al. 2013; Dodds et al. 2011; Horton et al. 2011; Keating et al. 2014).

In contrast, push strategies target people who are doing something else online other than actively looking to participate in research. One of the push methods is advertising on social media websites such as Facebook and Google+ to intercept social media users (e.g., Ramo and Prochaska

Table 1. Selected Types of Online Recruiting Sources.

Recruiting source	Definition	Recruiting flow	Cost model for researcher	Options for targeting recruits
Online classifieds	Sites where classified advertisements are posted (e.g., Craigslist, Backpage, and Adoos); recruiting advertisements are usually posted in the volunteer section	Pull	Free	None
Crowdsourcing	Websites that distribute a labor-intensive task that is not easily automated across a large pool of workers (e.g., Mechanical Turk and InnoCentive)	Pull	Pay per completed task	None
Search engine advertising	Services that display ads with search results and affiliated apps/sites (e.g., Google AdWords and Yahoo!)	Push	Pay when ads are clicked or shown	Can target based on user demographics and key words in a user's search
Social media	Networking sites that show ads on part of the screen (e.g., Facebook, Google+, Twitter, and Pinterest); Ads are displayed when user's profile fits the criteria set up by the advertiser	Push	Pay when ads are clicked or shown	Can target based on user demographics and interests

2012). Another push method is search engine advertising (e.g., Google and Yahoo), whereby researchers advertise to online users as they browse various web pages (e.g., Graham et al. 2006; Ramo et al. 2010). We make this push versus pull distinction (labeled "recruiting flow" in Table 1) because we speculate that the two methods will have important implications for the types of people recruited and the quality of data they provide. It is possible that volunteers recruited using a pull strategy (online classifieds and crowdsourcing) will be more motivated because they have the intention to work; it is also possible that they will be less motivated because they merely want to collect their reward before moving on to the next task. Early evidence suggests that volunteers recruited using a pull strategy do provide data of adequate quality. Paolacci et al. (2010), for example, found that results obtained from MTurk were similar to results obtained from other samples for a series of well-known psychology experiments. In addition, researchers have found the reliability of the data collected using MTurk participants to be comparable to those from other convenience samples (Behrend et al. 2011; Buhrmester et al. 2011).

Yet, other measures of data quality remain unexplored. For example, a pervasive characteristic of survey respondents is their tendency to satisfice—that is, failing to expend the amount of effort required to provide the most thoughtful answers (Krosnick 1991). To our knowledge, little research has investigated whether people recruited from push versus pull sources differ in the amount of effort they expend in a survey. Moreover, survey data can also be compromised if respondents are not willing to share their information, especially information they consider to be personal or sensitive. Hence, it is important to find out whether participants recruited from different resources differ in their disclosure of personal information.

There are other differences between these recruitment sources besides whether they are a push versus pull methods. In Table 1, we highlight two such differences, though there are certainly others. One difference is related to cost. Online classified ads are free; most social media and search engine advertising charge according to the numbers of times that ads are shown or clicked by the users; and crowdsourcing sites charge when a task is completed. Regardless of their fee structure, there is evidence that online recruiting methods tend to be more cost-efficient than traditional recruiting methods. For example, Gordon et al. (2006) recruited participants for a smoking cessation study and found an average cost per participant of US$6.70 on Google AdWords compared to US$36 for a direct mailing campaign and US$115 for print ads in a newspaper. Traugott (2012) recruited respondents for a political survey and found an average cost per participant of US$0.83 on MTurk, which was far less expensive than two volunteer online panels (Qualtrics: US$5.00 per case and Zoomerang: US$5.57 per case). Yet, these studies used only one new recruiting method, so little is known about how the different new methods compare to each another in terms of cost efficiency (e.g., the cost of Google AdWords vs. MTurk).

Another difference between online recruitment sources is related to targeting highly defined samples (e.g., smokers, young adults, and iPhone users). In online classified ads or crowdsourcing, researchers cannot generally control who sees their ads. On social media and search engines, however, researchers can target based on user demographics or interests when advertising. For example, in a study about young adult substance abuse, Ramo and Prochaska (2012) used a Facebook campaign to target a very specific subgroup: 18- to 25-year-olds with "tobacco- and/or marijuana-related key words that appeared in their Facebook profiles." Search engine advertising also offers a different type of targeting by allowing researchers to specify key words for their study. For instance, Morgan et al. (2013) chose "depression test" for a depression intervention study when using Google AdWords.

Given all these different ways of recruiting participants, an important question arises as to how to choose a recruitment method and whether the particular choice matters. In this article, we report our experience recruiting a convenience sample of US iPhone users using four different online sources: Craigslist, Google AdWords, Facebook, and MTurk. Specifically, we compare their cost efficiency, the demographics of the recruits, and the quality of data they provide (based on how they respond to a web survey). Our findings reveal important differences among the four recruitment sources that we believe can help researchers make more informed decisions on where to recruit their participants.

Methods

We recruited participants for a methodological study about the quality of survey data collected in voice and text message interviews on smartphones, in particular, iPhones (see Schober et al. 2015). All procedures were approved by the Health Sciences and Behavioral Sciences Institutional Review Boards at the University of Michigan. Recruits were asked to complete a 13-question web screener to determine eligibility for the methodological study. Our analysis focuses on respondent behavior in this web survey, not behavior during the follow-up interview on iPhones, so that we can directly compare the original samples coming from different sources. (The follow-up interview filtered out those who did not finish the screener questionnaire or did not meet our qualification criteria and were thus not invited to the main interview on their iPhone.) The recruiting period ran from March 13 to September 6, 2012, with a break from May 10 to July 16 because no interviews for the methodological experiment were conducted during that time.[1] In total, we recruited 3,280 recruits who were older than 21 years of age and provided a U.S. phone number associated with an iPhone.[2] The screener survey was programmed to capture where the participants came from, and we were able to identify the recruitment source for 2,885 participants.

Four Recruitment Methods

In this section, we describe how we used each of our four online recruiting sources. Note that in all of our advertisements, we included a URL link to our web screener.

Craigslist. Our advertisement was posted in the "volunteers" category of the "community" section in several major metropolitan areas (Detroit, Houston, Seattle, etc.). Since most posts generated only about one or two screened-in participants per day, we tried to boost recruitment by posting in more cities. However, not long after we adopted this strategy, Craigslist started removing the postings, sometimes with notification and sometimes without it.[3] We reverted back to our original strategy (of posting in one or two cities at a time) for the remainder of the recruitment campaign.

Crowdsourcing on MTurk. We posted several announcements, which Amazon calls human intelligence tasks (HITs), that workers could browse and decide whether or not to complete our web survey. In our HITs, we noted that the task was to complete our web survey (and that a sample of eligible participants would be selected for a longer interview on their iPhone with a separate payment upon completion). The wording of one of the HITs is shown in Online Appendix A. The task was restricted to U.S-based workers. After completing the screener but before clicking submit, workers were instructed to enter the last four digits of their iPhone number directly into the HIT's text field; we compared these numbers to what they entered in the screener (to verify that the person who submitted the HIT also completed our screener).

For our general HIT, we first offered a US$0.30 reward and then US$1.00 to speed up recruitment. Although MTurk does not provide explicit options for targeting workers based on their demographics, toward the end of our recruitment campaign we attempted to recruit older iPhone users by posting HITs with age requirements and offering a US$2.00 reward. One HIT stated that participants must be 45–54 to participate; another stated that participants must be 55 or older. This approach was slower in that it produced fewer recruits per day than the original HIT, but it did successfully attract older recruits. To prevent workers from registering for all of the HITs (each change in payment or addition of restrictions was treated as a separate HIT), we added instructions to complete our web screener only once, and we only approved payment the first time the HIT was completed by a particular worker.

Paid advertising on Google. We launched a Google AdWords recruitment campaign by creating a text ad and a list of key words that we defined, such as "iPhone." Given the space limitations—130 characters—we could only highlight what we deemed to be the most important information (see a screenshot of the ad in Online Appendix A). Note that we misstated the time required to complete the interview for the methodological study in this advertisement: It states that we are recruiting for a 20-minute survey not a 30-minute one like the other advertisements. This might have made recruiting on Google AdWords slightly easier than it would have been had we stated that it was a 30-minute interview. Thus, Google AdWords might have been even more expensive than what we observed here if the main interview length had been correctly stated as 30 minutes as in the other recruitment sources.

We set our bid price at US$0.85 per click, as suggested by Google for some of our key words. It turned out that the vast majority of clicks on our ad came from websites in Google's display network[4] such as the *New York Times* online, not the paid search results, likely because our bid price for key words like "iPhone" was less than other advertisers were willing to pay to appear in the search results. We budgeted US$50–US$300 per day. In total, our ads generated 61,407 unique clicks, leading to 827 eligible participants who completed the screener.[5]

About halfway through the recruitment process, we realized that our recruits tended to be young compared to the U.S. population, so we targeted older iPhone users by changing our ad settings to only show our ad to those users that Google identifies as 55 and older. While these targeted ads cost more per click (on average) than the general ads (US$0.45 vs. US$0.11), they were successful in recruiting older respondents.

Table 2. Cost, Duration, and Number of Recruits by Recruiting Method.

Recruiting method	Total cost	Duration of recruiting campaign (days)	Number of recruits	Recruits per day (number of recruits/ duration)	Average cost per recruit (total cost/number of recruits)
Mechanical Turk	US$2,929	72	1,750	24.3	US$1.67
Craigslist	None	110	299	2.7	None
Google AdWords	US$9,495	102	827	8.1	US$11.48
Facebook	US$272	11	9	0.8	US$30.22

Paid advertising on Facebook. We targeted those who live in the United States, are age 21 and older, "like" the iPhone, and speak English. This led to approximately 1.4 million users on Facebook (estimated by Facebook as of March 2012). A screenshot of the ad can be found in Online Appendix A.

We set the maximum price we would pay for each click on the ad at US$1.00; this was within the price range suggested by Facebook, given our target audience. The campaign generated 286 unique clicks on our ad and nine eligible participants who completed the screener. Because this ad campaign was expensive, we relied mostly on other recruitment methods.

Results

Cost and Efficiency

As Table 2 shows, in general, the pull methods were more efficient than the push methods. MTurk, a pull method, attracted the most total recruits (1,750) and the most recruits per day (24.3). Craigslist—another pull method—was most cost-efficient because it was free to post, and MTurk was second with an average cost of US$1.67 per recruit. The push methods, in contrast, were more expensive. On average, Google AdWords cost about US$11.48 and Facebook cost US$30.22 to recruit each participant. An important difference between the two types of methods is that MTurk participants were paid only when they completed the web survey, while Google AdWords and Facebook charged whenever someone clicked on our advertisement, even if they failed to complete the web survey. This inflated our recruiting costs because many people who clicked on the ad quit before seeing the first screener question. Thus, although the cost-per-click charge was small for Google AdWords and Facebook (US$0.15 and US$ 0.95, respectively), the resulting cost per screened-in participant was actually quite high compared to MTurk and Craigslist.

Demographics of Recruited Participants

Although this study did not attempt to obtain a representative sample of iPhone users, we were still quite interested in how participants recruited from different methods would compare demographically.[6] Specifically, we focus on the comparisons of gender, age, and income (see Online Appendix B, Figures B1–B3, respectively), because there are published data on these characteristics of U.S. national iPhone users[7] that we can use as benchmarks. Because only nine participants were from Facebook, the comparisons of the demographics and behaviors in the rest of the results section focus on MTurk, Craigslist, and Google AdWords.

As Figure B1 shows, Google AdWords yielded a mix of female and male iPhone users very close to the national estimates of iPhone users at the time of the study (47% female and 53% male). However, both of the pull methods overrepresented female users (MTurk: 52% and Craigslist: 68%).

As mentioned earlier, about halfway through our study we began to target older adults in the ads on MTurk and Google to diversify the age of our participants. To make fair comparisons of participants' age from different recruitment sources, we excluded those recruited through the age-targeted advertisements. Figure B2 presents the findings when the ads only stated "21 years or older" in all three recruitment methods. The national benchmark data show that iPhone users during this period tended to be younger than the full population, with about 10% of users between 55 and 64 years old and 8% aged 65 or older. Our recruits from Google AdWords, although still young, were more diverse and more similar to the national iPhone users with respect to age than the recruits from the other sources; MTurk and Craigslist yielded almost no recruits in the 65 and older age-group.

As Figure B3 shows, the benchmark data also indicate that the majority of U.S. iPhone users in this period were affluent—60% of all iPhone users had a household income of US$75,000 or more. However, those earning US$75,000 were underrepresented in all our recruitment methods, among which Google AdWords seemed to recruit more affluent participants than MTurk or Craigslist.

In summary, the push method—Google AdWords—attracted a sample with a demographic distribution that was more similar to the target population than either of the pull methods (Craigslist and MTurk).

Reporting Behavior of Recruited Participants

In addition to the demographics, we are interested in whether participants recruited from different methods behaved differently when filling out our screener questionnaire. We first examined how long it took people to complete the online screener. On average, Google AdWords recruits spent about twice as long as those from Craigslist and MTurk in completing the screener (Google AdWords: 254 seconds; MTurk: 128 seconds; Craigslist: 153 seconds; and $F(2, 4{,}576) = 1{,}301.57$, $p < .001$). The same pattern is observed after controlling for age and education (Online Appendix C, Table C1).

Why is there such difference in completion times across recruitment methods? The difference holds across all age-groups, so it is unlikely a result of the age differences across the different recruitment methods. One possibility is that longer times of Google AdWords participants suggest they are more conscientious respondents (i.e., they expended more effort answering the questions to provide better quality data). To examine this, we looked at three indicators of insufficient reporting effort: incomplete answers to a zone improvement plan (ZIP) code question and *don't know* responses to the two questions asking about participants' cell phone voice plan and text plan. In our web survey, Google AdWords participants were more likely to choose *don't know* to questions about their cell phone voice plan (Google: 17.4%; MTurk: 15.8%; Craigslist: 12.3%; $\chi^2 = 7.94$, $df = 2$, and $p = .019$) and text plan (Google: 9.7%; MTurk: 6.4%; Craigslist: 5.3%; $\chi^2 = 23.37$, $df = 2$, and $p < .001$). Google AdWords participants also provided more incomplete ZIP codes than recruits from the two other methods (Google: 23.9%; MTurk: 9.0%; Craigslist 2.3%; $\chi^2 = 537.49$, $df = 2$, and $p < .001$), even after controlling for demographic characteristics including age, gender, education, and income levels (Online Appendix D, Table D1). So, it does not appear that the longer completion times of Google AdWords recruits result from more effort in answering the questions.

What caused the longer completion times of Google AdWords recruits? The additional analysis suggests that this might be because Google AdWords participants were more cautious about giving out personal information. We found that participants from Google AdWords were more likely to not provide their cell phone number (Google: 32.7%; MTurk: 22.8%; Craigslist 22.2%; $\chi^2 = 83.62$, $df = 2$, and $p < .001$) and to not provide their income (Google: 18.9%; MTurk: 6.3%; Craigslist: 8.6%; $\chi^2 = 290.06$, $df = 2$, and $p < .001$) compared to those from MTurk and Craigslist. This difference remains significant even after we controlled for participants' demographics. (Detailed regression results are presented in the last two columns of Table D1.)

Thus, the recruits from the pull methods (Craigslist and MTurk) provided better data (i.e., fewer *don't know* responses, fewer incomplete ZIP code answers, and more disclosure) than the recruits from the push method (Google AdWords).

We wondered whether recruits also behaved differently in the follow-up interview on people's iPhones, even though many participants who started the screener did not make it to the main study interview (which is why this is not our focus). We found no evidence of data quality differences in the main interview between Google AdWords and the other two sources: average interview length, $F(2, 1{,}157) = 0.19$, and $p =$ n.s.; number of rounded answers, $F(2, 1{,}159) = 1.07$, and $p =$ n.s.; and number of socially undesirable answers to sensitive questions $F(2, 1{,}159) = 0.99$, and $p =$ n.s. In other words, the choice of recruitment source had an impact on data quality in the screener questionnaire but not in the follow-up study. One possible explanation for why the effect disappeared is that the pull method recruits who were most reluctant to provide thoughtful answers or provide personal information dropped out before making it to the main interview. It may well be that differences in data quality would have emerged in the main interview had we not asked these particular screener questions. If this is the case, then our observed effect of recruitment source on data quality may generalize not only to screener surveys but to research studies that do not use a screener.

Discussion

The results indicate very different performance between two different recruitment strategies: those that pull in volunteers actively looking for opportunities (MTurk, Craigslist) and those that push out a recruiting ad to intercept online users who are engaged in other unrelated online activities (Google AdWords and Facebook). In terms of cost efficiency and data quality, the pull methods performed best. For cost, Craigslist was free, and MTurk cost less than one-sixth as much per recruit as Google AdWords and less than one-fifteenth as much per recruit as Facebook. We think this is because the push ads had to capture people's attention and persuade them to switch from what they were currently doing in our online survey. These ads presumably needed to have positive attributes (interesting topic and cash incentive) that were made salient by the ad (Groves et al. 2000). The pull ads, in contrast, only had to bring in those recruits who were already looking for paid work, so seeking their participation tended to be easier.

For data quality, the pull method recruits (Mturk and Craigslist) provided more thoughtful answers and were more willing to provide all of the necessary information to be recruited compared to their push method counterparts (Google). Specifically, the pull method recruits provided fewer *don't know* responses to the questions about their cell phone usage and fewer incomplete ZIP codes than the push participants. They were also more likely to provide their cell phone number and report their household income. In short, those who are actively looking for online work appeared to be more committed to the survey task than those intercepted through online advertisements. MTurk may also have stood apart from the other three methods to the extent that workers believed that they would not be paid unless they completed the entire task, answering all survey questions.

While the pull methods performed best in terms of cost efficiency and data quality, the push method (i.e., Google AdWords) attracted a more diverse pool of recruits. Specifically, Google AdWords attracted recruits who were relatively close to our target population in terms of their age, gender, and income. Meanwhile, the pull methods yielded samples that skewed young, female, and low income. It is possible that the demographic differences are not inherent in the distinction between push and pull strategies. For example, although the two pull strategies in this study (Mturk and Craigslist) led to pools of recruits that heavily skewed young, it is easy to conceive of the opposite when a pull strategy is used to draw participants from someplace else, such as a pool of retirees looking for work.

We should note two limitations of this study. First, it is unknown how the specific features of our study affected recruitment, and we acknowledge that our results may depend on some study-specific factors, such as eligibility criteria (i.e., iPhone users), the wording of our advertisements, the maximum price-per-click chosen on Google AdWords and Facebook, and the size of the MTurk reward that was offered. Second, we compared only two push methods and two pull methods. We acknowledge that the results could change if the study population were different (e.g., smokers and Spanish-only speakers), and if other recruiting methods were used, such as online web panels. Although paying an online panel to invite their volunteers to participate in a survey would also fall under the pull category (e.g., Chang and Krosnick 2002; Yeager et al. 2011), we didn't explore this method in our study mostly because it appeared to be less straightforward compared to other methods for which we could manage the advertisements directly. We suspect that the recruitment performance of online panels is likely to be similar to the two pull methods examined in our study (i.e., recruits would be more cost efficient and committed to the survey task than those recruited from push methods such as Google AdWords). Further studies can further test the push versus pull distinction by expanding the set of recruitment methods examined (e.g., Google+, river sampling strategies that use pop-up ads, or banner ads to intercept online users).

What are the practical implications of these findings? Since no one recruiting source was superior in every measure, there is no one-size-fits-all recommendation and it is worth exploring different methods to find an approach that suits a study's purpose and budget. If a researcher is looking for a low-cost method, then a pull strategy like MTurk or Craigslist appears to be superior. Likewise, if a study demands substantial commitment from participants (e.g., sensitive questions and complicated tasks), then a pull strategy is perhaps more appropriate. However, as this study shows, a push strategy like Google AdWords is sometimes needed to recruit a more diverse sample. It should also be noted that we are not advocating for the use of convenience samples in all cases as there are certainly instances when a probability-based sample is better suited for a study's purpose.

The findings of this study suggest that pull versus push could be a defining distinction for different recruitment strategies for convenience samples. Given that online recruitment sources are rapidly evolving and expanding in number, we encourage more research to test this push versus pull framework for categorizing online recruiting sources. More generally, we propose that this kind of organizing framework, and probably new theoretical distinctions, will be important for guiding the choice of researchers interested in recruiting online; given the rapid pace of change in technologies and online behavior, theories and frameworks will be needed that endure beyond particular recruiting sources.

Acknowledgments

The authors thank Andrew Hupp, Huiying Yan, Lucas Vickers, Patrick Ehlen, Stefanie Fail, Lloyd Hemingway, Michael Johnston, Kelly Nichols, and Leif Percifield for their help with this study. The authors also thank two anonymous reviewers as well as Florian Keusch for helpful comments on earlier drafts of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the

National Science Foundation grants [SES-1026225 and SES-1025645] to F.G.C. and M.F.S.

Notes

1. The exact dates for each method are as follows: Craigslist (3/13–5/9; 7/1–9/6), MTurk (4/20–5/9; 7/17–9/6), Google AdWords (3/21–5/9; 7/17–9/6), and Facebook (3/16–3/18; 8/17–8/24).
2. As the last part of the screen-in process, we sent participants a text message containing a link to a web page. Participants were instructed to tap the link to confirm their eligibility to participate in this study. When they accessed the web page through the link, we captured the user-agent string to determine whether their device was actually an iPhone.
3. This is often referred as "Craigslist Ghosting" when the user gets a confirmation of the ad successfully posted, but the ad is not actually shown on Craigslist. It occurs as a result of Craigslist's attempt to reduce "spamming."
4. "Display network" is the name for a group of thousands of websites, including news sites and blogs, where Google's ads can appear. Google shows ads on the display network by default, and we did not change this setting.
5. Because we wanted to target iPhone users, we specified in the Google AdWords settings our ad to be shown on smartphones with iOS operating systems (i.e., iPhones). We briefly opened up our ad to show on tablets and computers as well, in an attempt to reach a broader audience. Unexpectedly, this approach led to fewer screened-in participants and a higher cost per click. Thus, we decided to return to the original approach of targeting iOS smartphones exclusively.
6. Diversity with respect to demographics might be important to a researcher who wants to estimate treatment effects across subgroups or use weighting adjustments to make inference to a larger population.
7. Radwanick (2012) used data derived from a nationally representative sample of mobile subscribers to report the demographic profile of iPhone users from March to May 2012, which is the same time we began recruiting iPhone users.

Supplemental Material

The online data supplements are available at http://fmx.sagepub.com/supplemental.

References

Anderson, S., S. Wandersee, A. Arcenas, and L. Baumgartner. 2013. Craigslist samples of convenience: Recruiting hard-to-reach populations. http://fiesta.bren.ucsb.edu/~sanderson/CraigslistSurvey.pdf (accessed October 25, 2013).

Baker, R., M. J. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. J. Gile, and R. Tourangeau. 2013. Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology* 1: 90–143.

Behrend, T. S., D. J. Sharek, A. W. Meade, and E. N. Wiebe. 2011. The viability of crowdsourcing for survey research. *Behavior Research Methods* 43:800–13.

Berinsky, A. J., G. A. Huber, and G. S. Lenz. 2012. Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis* 20:351–368.

Buhrmester, M., T. Kwang, and S. D. Gosling. 2011. Amazon Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* 6:3–5.

Chang, L.C., and J. Krosnick. 2002. A comparison of the random digit dialing telephone survey methodology with Internet survey methodology as implemented by Knowledge Networks and Harris Interactive. Paper presented at the 98th Annual Meeting of the American Political Science Association, Boston, MA, August 28–September 1.

Crump, M. J. C., J. V. McDonnell, and T. M. Gureckis. 2013. Evaluating Amazon Mechanical Turk as a tool for experimental behavioral research. *PLOS ONE* 8: e57410.

Dodds, P. S., K. D. Harris, I. M. Kloumann, C. A. Bliss, and C. M. Danforth. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLOS ONE* 1:6.

Gordon, J. S., L. Akers, H. H. Severson, B. G. Danaher, and S. M. Boles. 2006. Successful participant recruitment strategies for an online smokeless tobacco cessation program. *Nicotine & Tobacco Research* 8:S35–41.

Graham, A. L., B. C. Bock, N. K. Cobb, R. Niaura, and D. B. Abrams. 2006. Characteristics of smokers reached and recruited to an internet smoking cessation trial: A case of denominators. *Nicotine & Tobacco Research* 8:S43–48.

Groves, R. M., E. Singer, and A. Corning. 2000. Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly* 64: 299–308.

Henry, G. T. 1990. *Practical Sampling*. Vol. 21. Newbury Park, CA: Sage.

Horton, J. J., D. G. Rand, and R. J. Zeckhauser. 2011. The online laboratory: Conducting experiments in a real labor market. *Experimental Economics* 14: 399–425.

Keating, M., B. Rhodes, and A. Richards. 2014. Crowdsourcing: A flexible method for innovation, data collection, and analysis in social science research. In *Social Media, Sociality, and Survey Research*, eds. C. A. Hill, E. Dean, and J. Murphy, 179–210. Hoboken, NJ: Wiley.

Krosnick, J. A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5:213–36.

Morgan, A. J., A. F. Jorm, and A. J. Mackinnon. 2013. Internet-based recruitment to a depression prevention intervention: Lessons from the Mood Memos study. *Journal of Medical Internet Research* 15:e31.

Paolacci, G., J. Chandler, and P. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5:411–19.

Radwanick, S. 2012. 5 years later: A look back at the rise of the iPhone. ComScore Blog. http://www.comscore.com/ita/Insights/Blog/5_Years_Later_A_Look_Back_at_the_Rise_of_the_iPhone (accessed January 13, 2014).

Ramo, D. E., S. M. Hall, and J. J. Prochaska. 2010. Reaching young adult smokers through the Internet: Comparison of three recruitment mechanisms. *Nicotine & Tobacco Research* 12:768–75.

Ramo, D. E., and J. J. Prochaska. 2012. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. *Journal of Medical Internet Research* 14:e28.

Schober, M. F., F. G. Conrad, C. Antoun, P. Ehlen, S. Fail, Andrew L. Hupp, Michael Johnston, Lucas Vickers, H. Yanna Yan, and Chan Zhang. 2015. Precision and disclosure in text and voice interviews on smartphones. *PLOS ONE* 10: e0128337.

Traugott, M. W. 2012. Data quality from low cost data collection methodologies. Paper presented at the 67th annual conference of the American Association for Public Opinion Research, Orlando, FL, May 17.

Worthen, M. 2013. An invitation to use Craigslist ads to recruit respondents from stigmatized groups for qualitative interviews. *Qualitative Research*. http://qrj.sagepub.com/content/early/2013/04/02/1468794113481791 (accessed January 13, 2014).

Yeager, D. S., J. A. Krosnick, L. C. Chang, H. S. Javitz, M. S. Levendusky, Alberto Simpser, and Rui Wang. 2011. Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly* 75:709–47.