**Best Practices for Working with Opt-In Online Panels**

Jason Peifer, with minor additions by Kelly Garrett

**Abstract:**
This manuscript provides a brief introduction to opt-in online panels and offers a comprehensive list of procedural principles that can help to improve data quality. The guidelines are organized according to the topics of recruitment, survey instrument quality, identifying careless responses, minimizing dropouts, and preparation for post-survey procedures.

Amid the continued expansion of Internet access, strong demand for consumer data in commercial market research, the relatively low costs of Internet-based research, and declining response rates of more traditional phone-based survey research (Pew Research Center, 2012), the use of online surveys/experiments has become popular over the last fifteen years (Baker et al., 2010). However, although online surveys function as a powerful research tool, they are replete with methodological challenges. In particular, many online surveys/experiments that use volunteer, “opt-in panels” face unique hurdles compared to more traditional probability-based samples and the convenience samples of a university laboratory setting.

Opt-in panels consist of individuals who are willing to join a pool of online survey participants, usually in return for compensation like reward points, coupons, being entered in a drawing, or cash. Such participants might be recruited by a banner ad on a website, direct mail, email (Baker et al., 2010), or by word of mouth. Companies that provide opt-in online panels include Qualtrics Panels, Survey Sampling International (SSI), Toluna (formerly Greenfield), and uSamp. There are also a few companies that use variations on opt-in panels. For example, YouGov (formerly Polimetrix) offers what it calls a “matched sample,” using census data to guide the selection of participants from a large opt-in panel, and constructing weights that the company claims can be used to approximate a representative sample (Rivers, & Bailey, 2009).

Last revised: April 25, 2014
Opt-In Online Panel Best Practices

Amazon’s Mechanical Turk (MTurk) also relies on a large online panel of volunteers who complete short tasks, including surveys and experiments, for micropayments that Amazon administers. Although not originally designed as a research tool, MTurk has been used for these purposes for the past several years (Berinsky, Huber, & Lenz, 2012).¹

The key characteristic of “opt-in panels” is that the participant pool is not constructed with random selection. Rather, the group of participants is comprised of self-selected individuals who choose to sign up with a panel, participating at will. A probability-based sample, in contrast, is comprised of subjects who are randomly selected by a researcher/survey company, in which everyone in the target population theoretically has a non-zero chance of being selected (Groves et al., 2009). GfK (formerly KnowledgeNetworks) is a well-known example of a company utilizing a probability-based sample for online data collection. Probability sampling helps to bolster confidence that that sample is representative of its target population. In contrast, the representativeness of a non-random sample is always questionable.

In an effort to aid the collection of high quality data via opt-in panels, below I offer a comprehensive list of various procedural principles to implement—or at least consider—when preparing to field an online survey/experiment. The guidelines are organized according to the topics of recruitment, survey instrument quality, identifying careless responses, minimizing dropouts, and preparation for post-survey procedures.

Recruitment

The recruitment aspect of the data collection process broadly relates to addressing

¹ MTurk has a number of unique limitations, including a highly skewed sample (e.g., disproportionately male, well educated, and liberal) and the presence of discussion boards on which member of the panel discuss projects and project sponsors.
potential issues of coverage error, ensuring that the sample includes the types of participants necessary in order to be confident about the applicability of the estimates provided by the data.

**Identify target population.** An important initial step when designing a survey is to explicitly identity the target population of the study’s investigation (Groves et al., 2009, p. 30). In this light, the researcher can assess whether a given opt-in panel can adequately provide a sample that matches the target population. Depending on the nature of the study’s goals and the key variables in focus, it may not be imperative to have a nationally representative sample. Yet it is usually advantageous to have a more diverse sample than one might find in a university student research pool. Most panel companies, including Qualtrics Panels and SSI, allow researchers to set up demographic quotas in order to ensure sample diversity. The result is still a convenience sample—it is *not* representative—but its heterogeneity may be useful.

**Consider a dual sample.** Depending on the goals of the study, it may be worth employing a dual sample. That is, augmenting panel survey data with student sample data. Combining a student sample with a non-student sample (i.e., an online opt-in panel) not only provides a study with more power, but also allows for comparison between the two populations such that the researcher can better assess whether the findings are an artifact of each unique sample or are due to a more generalizable phenomenon. Furthermore, as Druckman and Kam (2011) explain, “when dual samples are not feasible, researchers can take a second-best approach by utilizing question wordings that match those in general surveys (thereby facilitating comparisons)” (p. 54).

**Leverage institutional credibility.** When crafting a survey, it may be helpful to note one’s institutional affiliation in the survey. Fan and Yan (2010) report that surveys sponsored by
academic and governmental agencies tend to have higher response rates than commercial surveys. By mentioning that the survey is associated with, for example, The Ohio State University, the researcher may be able to leverage the university’s credibility. Of course, academia is not universally credible, so this tactic should be thoughtfully considered, depending on the sample that the researcher intends to use.

Guard against multiple submissions. Whether using an opt-in panel or a student sample, it is important to safeguard against multiple submissions—both intentional and accidental multiple submissions. As Fan and Yan (2010) suggest, the researcher might consider a semiautomatic login procedure in which a unique username or a password is assigned to each participant. (By using an automatic login, in which a unique identifier is embedded into the URL, the respondents will not need to enter a password for access.) In this way, the researcher can better ensure that a given person’s data is only used once. Many panel companies (including Qualtrics Panels and SSI) will manage this part of the process.

Survey Instrument Quality

When crafting and compiling the question items to be used in an online survey, it is important to ensure that the survey items will be clearly and quickly understood, so as to avoid measurement error (due to the researcher). There are a variety of simple but easy-to-overlook pointers worthy of consideration when designing any survey.

Double-barreled questions. A survey item should not pose two questions at once. The use of such questions casts doubt on the reliability of the survey, as there is uncertainty about which question—in the double-barreled inquiry—was being answered by survey’s respondents.

Negatively worded questions. The use of negative words in a survey item can create
unnecessary confusion, especially when an answer option is itself a negative response (e.g., “Do you agree or disagree with the argument that abortion should not be outlawed?”).

**Context effects.** Researchers should be cognizant of how the placement of a question might inadvertently affect the interpretation/engagement with subsequent survey items. Randomizing the question ordering can sometimes help in this regard. For instance, in terms of a single index (e.g., need for cognition), it usually advisable to randomize the order of the various questions—an easy task when administering a survey online.

**Wording effects.** This variation of measurement error refers to how certain question wording can create “leading questions,” thus biasing participant responses in unintended ways. In this regard important to use neutral wording. Furthermore, it is important to ensure that the wording used has a relatively unambiguous, consistent meaning for virtually everyone. For instance, asking how much one appreciates “political entertainment” could be problematic because this term could easily connote different things for different people.

**Aid memory recall.** It can be helpful to build retrieval cues into survey questions to help jog people’s memory about behaviors and events (Groves et al., 2009). For instance, if inquiring about media exposure, it may be helpful to offer a concrete time frame from which to produce an estimate, rather than an abstract “on average” type of time frame (e.g., use “in the last week” rather than “in typical week”).

**Consistent use of scales.** Researchers should strive to be consistent with the scales used in a survey. Needlessly changing the format to the question response options can lead to measurement error, in which respondents erroneously use the response options of earlier questions to answer subsequent questions. Notably, Groves et al. (2009) advise using 5- or 7-
point scales for attitudinal measures, starting with the end of scale that is least popular.

**“Loading” sensitive questions.** If a question pertains to sensitive information that, perhaps for reasons of social desirability, individuals may be reluctant to answer honestly, it may be worthwhile “loading” the question in order to reduce misreporting (Groves et al., 2009, p. 246). Loading a question means to word the survey item in such a way that encourages the socially undesirable answer. For instance, the a question might begin with, “Many people smoke marijuana now and then…” By implicitly suggesting that such behavior is not abnormal, this preface may prompt more honest responses.

**Grouping questions by topic.** It is often helpful to group together questions on the same topic (e.g., the items of an index), as opposed to loading a single item on each page. This practice enables respondents to complete the survey more efficiently. However, this is not to suggest that all questions should be listed on one page. Instead, a paging design that limits a page to a handful of questions (thus, not needing to scroll down the page) is preferable. Moreover, a paging design provides the researcher with greater control over the order of the questions and allows for the use of automated skipping—so as to avoid subjecting respondents to irrelevant questions (Tourangeau, Conrad, & Couper, 2013).

**Pilot survey.** It is good practice to pilot a web survey with a small group of participants (e.g., friends, relatives, and/or students) before activating the survey. This enables the researcher to identity and revise problematic questions and survey formatting (Fan & Yan, 2010).

**Identify Careless Responses**

Due to the self-administered nature of web surveys, measurement error due to the respondent is a common problem in online opt-in panel survey data. In particular “satisficing”—
which involves a respondent exerting minimal cognitive effort in order to quickly complete the survey—erodes the quality of survey data. Therefore, a quality online survey should be designed in such a way that (a) satisficing is minimized (b) items are included that provide the capacity to identify careless responses.

**Trap questions.** The use of “trap questions” can help researchers gauge the quality of their online panel (Gittelman & Trimarchi, 2012). Trap questions help to flag respondents who are not paying close attention to directions and not providing consistent responses. For instance, a researcher might employ a quality check trap question like, “To show that you are reading these instructions, please leave this question blank” or “For data quality purposes, please select 4.” Meade and Craig (2012) suggest incorporating these sort of questions every 50 to 100 survey items, up to a maximum of three times.

**Manipulation check of stimulus:** Incorporate manipulation checks to ensure that participants devoted sufficient attention to the stimulus (if conducting an experiment). For example, after participants view, say, a video embedded in the survey-experiment, the survey might ask the viewer identify from a list (1) which program was just shown and (2) the central topic of the video clip. In this way, the researcher can be more confident that any resulting effects are (or are not) due to the experimental stimulus.

**Self-report of study engagement:** Consider posing a question at end of survey that asks participants to self-report the quality of the data they provide. For example, a survey might ask: “Lastly, it is vital to our study that we only include responses from people who devoted their full attention to this study. In your honest opinion, should we use your data?” (see Meade & Craig, 2012). Importantly, if a student sample is involved, it should be emphasized that they will
receive credit for their participation regardless of how they respond to such a question.

**Record response time.** It is important to record the length of time it takes each participant to complete the survey. This will enable the researcher to identify blatant “speeders” and be in a position to decide how to handle such data.

**Minimize Dropouts**

In addition to identifying careless responses, it is important to minimize non-response error—i.e., survey error related to missing data. This can be especially important if the questionnaire includes random assignment to an experimental condition. Assuming that some individuals will be unmotivated to complete the study for reasons beyond the survey itself, it is preferable to have such participants drop out early on—before administering the stimulus and the measurement of other key variables. (Notably, Reips [2002] finds that most dropout takes place at the start of a survey.) In other words, it is helpful to weed out the uncommitted participants sooner than later. This allows the researcher to better gauge when dropouts are a systematic artifact of experiment’s independent variable manipulation—for instance, dropping out because of a boring control condition—and not due to random uncommitted participants (Reips, 2002).

**Seriousness check.** One might consider emphasizing the importance of taking the study seriously, reminding participants that good science needs good data. The survey could pose a question that inquires about the likelihood of the participant completing the whole experiment. Reips (2002) found that if an online survey states that only data sets from respondents with a motivation for serious participation will be analyzed, the dropout rate is usually much lower.

**Warm up technique.** In order to keep dropout low during the critical parts of the survey, a “warmup technique” might be used, wherein the actual start of the survey is positioned several
pages deep into the online study (Reips, 2002). If feasible, questions of secondary or even less importance could be placed on the initial pages. Here again, the hope is to shed the uncommitted participants early on, rather than lose them late in the survey.

**Survey length.** Even as the use of quality check questions can enhance survey data, it is important to strike a balance by safeguarding the efficiency of the survey, in order to avoid respondents becoming tired of taking the survey—i.e., “respondent fatigue” (Ben-Nun, 2008). This involves determining whether the survey is asking only the essential questions and perhaps posing the cognitively taxing questions early on, when respondents are most focused, and saving simple demographic and political orientation-type questions for the end.

**Positioning of sensitive questions.** Consider posing survey items of a sensitive nature (e.g., income) toward end of survey. Ideally, the researcher will have built some trust and a sense of commitment with participant by this point, thus increasing the chances of recording a response. Similarly, it can be beneficial to announce any motivationally adverse factors at the beginning of the survey. For example, note upfront how long the survey should take (e.g., “This survey should take about 20 minutes to complete”). Such forthrightness may not only engender trust, but also weed out non-serious respondents.

**Preparation for Post-survey Procedures**

Throughout the process of designing an online survey and collecting data, there are several practices that a survey researcher should be prepared carry out so that other scholars are better able to assess the quality of the data.

**Identify flat-liners:** If a panel respondent’s responses exhibit little variance due to “flat-lining,” it may be worth dropping the respondent’s answers from data set. There are statistical
Opt-In Online Panel Best Practices

procedures that a researcher can consider employing in order to identify flat-liners (e.g., see Menictas, Wang, & Fine, 2011).

**Report participation rate.** Even as survey panel companies are known to be reluctant to provide completion rate data, it is important to secure the participation rates of active online panel members (Menictas, Wang, & Fine, 2011, p. 36). Reporting a completion rate lends credibility to the survey, reflecting the quality of the recruitment procedures for participation in the survey. Importantly, participation/completion rates should not be confused with “response rates”—which are often not possible to calculate in a straightforward way, as with probability-based samples (Baker et al., 2013). Furthermore, an opt-in panel survey study should not report a margin of error (i.e., sampling error), as there is not a theoretical basis for sampling error with opt-in survey data (Weisberg, 2005).

**Full disclosure / Transparency.** Perhaps most importantly, researchers who use online opt-in panels should be forthcoming about the methods and strategies used to conduct their surveys. They should acknowledge the shortcomings of opt-in panels, not obscuring the weaknesses of non-random opt-in panel survey data. Not only does a full accounting of the data collection procedures allow other scholars to assess the quality of the data, it also enables others to replicate the study if so desired. To this end, AAPOR’s recommended wording for online opt-in surveys provides a helpful template to reference when reporting the study analyses. It reads:

> Respondents for this survey were selected from among those who have [volunteered to participate/registered to participate in (company name) online surveys and polls]. The data (have been/have not been) weighted to reflect the demographic composition of (target population). Because the sample is based on those who initially self-selected for
participation [in the panel] rather than a probability sample, no estimates of sampling error can be calculated. (AAPOR, n.d.)
Opt-In Online Panel Best Practices

References


Last revised: April 25, 2014
Opt-In Online Panel Best Practices


