



# Best Practices for Survey Recruitment

---

BRADY T. WEST, PH.D.

RESEARCH PROFESSOR

SURVEY RESEARCH CENTER, INSTITUTE FOR SOCIAL RESEARCH

UNIVERSITY OF MICHIGAN-ANN ARBOR

[BWEST@UMICH.EDU](mailto:BWEST@UMICH.EDU)

# 0. Screening vs. Main

---

A discussion of best practices in survey recruitment needs to start with the important difference between screening of sampled units (usually to determine eligibility for a survey) and the main data collection protocols for eligible sampled units who are selected for the main study

**Screening tips:** use small pre-paid incentives (e.g., \$2 bill); try inviting sampled units to screen via web first (if possible); be as inclusive as possible when collecting household rosters (e.g., nicknames); follow up initial non-respondents with priority mail and offers of larger incentives; try to leverage purchased commercial data to over-sample likely-eligible households (**West et al., 2015**)

**Main tips:** leverage information collected from the screener (if applicable) to predict key variables and response propensity; consider telephone *reminders* (not telephone interviewing; **West et al., 2023**)

**The remainder of this presentation will talk about additional best practices that apply to both stages!**

# 1. Use a Mix of Modes

---

To the extent possible (given project resources), try to **diversify and tailor contact strategies**, rather than continuing to use the same data collection mode over and over (e.g., making 30+ phone calls to the same number hoping to eventually make contact; **Dillman et al., 2014**)

Different individuals may be more receptive to certain modes (or, certain modes may not be possible, e.g., someone who does not have internet access)

**Fairly common sequential mixed-mode approach:** start with a mailed invitation to respond via web (cheapest mode, “push to web”); follow-up with a mailed paper option and a web reminder (still pretty cheap); maybe one more mailed invitation, followed by telephone *reminder* (if possible; **West et al. 2023**); follow up with in-person reminder / interview, if possible (most expensive)

## 2. Adaptive Survey Design (ASD)

---

Populations (and especially hard-to-reach populations) can be quite heterogeneous, and different aspects of survey invitations, such as the topic / sponsor / incentive / etc., may appeal differently to different people (**leverage-salience theory; Groves et al., 2000**); trying to use the same strategies for all different subgroups may produce bias and coverage issues!

**Adaptive Survey Design:** Divide your sample into groups based on auxiliary information that is available at the onset of data collection, and develop specialized protocols (different starting modes, different incentives, different contact strategies, etc.) for each group, based on historical analyses of the most effective recruitment strategies for each group (**Schouten et al., 2017**; see **Groves et al., 2009**, for a summary of such strategies for different settings)

Relies heavily on analyses of historical data, and being informed about what works best for the population that you are trying to study; review of past literature very important!

# 3. Responsive Survey Design (RSD)

---

Based on planned monitoring of key indicators of survey costs and survey errors (**paradata**) while a data collection is ongoing, determine if a given phase of the data collection has reached “capacity” and a new phase with a new protocol is likely to help (**Groves and Heeringa, 2006**)

Potentially apply a newer, more expensive protocol to a subsample of active cases in a later phase, and then weight respondents to account for the additional subsampling (**two-phase sampling idea**); introduces cost efficiency

Can embed experiments in earlier phases of data collection in an effort to determine what works best for what subgroups (see also #6 forthcoming); however, timing of interventions and protocol changes is crucial, and needs to recognize the past effort applied to a case

See **Axinn et al. (2021)** for a promising example of RSD using in-person follow-up in a smaller area

# 4. Case Prioritization

---

**Idea:** Shift effort (e.g., more call attempts) and resources (e.g., higher incentives) to active cases that are under-represented in the current respondent data set (to address nonresponse bias) or cases that are more likely to achieve yield goals (higher response propensity)

Other active cases still get worked, just not to the same extent

## Example Success Stories:

Peytchev et al. (2010)

Wagner et al. (2012)

West et al. (2021)

# 5. Stopping Rules / Limiting Effort

---

An idea borrowed from the clinical trials literature, which basically says to stop data collection effort on cases that are no longer likely to be fruitful

Shares similarities with case prioritization, but no longer applies effort to stopped cases

**Best candidates for stopping:** Cases that are 1) unlikely to reduce bias in multiple estimates (based on predictions of key variables), and 2) likely to increase costs substantially

**Rely heavily on good models of both key variables and expected costs;** careful modeling effort for these outcomes based on available data is absolutely critical

Good example of the trade-off between cost savings and bias / variance reduction

See Wagner et al. (forthcoming), Wagner and Raghunathan (2010), Rao et al. (2008)

# 6. Replicate Designs

---

**Idea:** Break a random sample up into random subsamples (replicates)

Release the random subsamples sequentially and not all at the same time (may be some overlap)

In earlier subsamples, design randomized experiments to test different data collection protocols

Build in “analysis time” in-between replicates to analyze the results of the experiments

Apply knowledge gained from earlier replicates to protocols used in later replicates to optimize data collection efficiency in those later replicates

Combine replicates into a single data set for analysis

**Examples:** Zhang et al. (2023), Wagner et al. (2012)



# 7. Hard-to-Reach Populations

---

Recruiting individuals from hard-to-reach (H2R) populations (e.g., low-income households, injection drug users, etc.) can be a very challenging task; often, sampling frames for such populations simply don't exist (coverage error!)

One option: **respondent-driven sampling (RDS)**, where you start with data collection “seeds” (e.g., patients in a clinic, participants in an aid program, etc.) and encourage those seeds to introduce the survey to other members of their networks (forming “chains” and ultimately producing a sample from the target population); see **Heckathorn (1997)** and **Goel and Salganik (2010)**

**Some great resources on the H2R topic:**

- \* A special issue of the Journal of Official Statistics (see **Willis et al., 2014**)
- \* An edited volume from a conference dedicated to the H2R topic (see **Tourangeau et al., 2014**)
- \* See also **Sosenko and Bramley (2022)** for recent work on smartphone-based RDS approaches
- \* Finally, **community-based recruitment approaches are crucial**, and QR codes may help (**Boelter et al., 2023**)

# References

---

Axinn, W. G., Wagner, J., Couper, M., & Crawford, S. (2021). Applying responsive survey design to small-scale surveys: Campus surveys of sexual misconduct. *Sociological Methods & Research*, 00491241211031270.

Boelter, Jacob, Alexis M. Dennis, Lisa K. Vogel, and Kenneth D. Croes. 2023. "Recruiting Hard-to-Reach Populations Amid the COVID-19 Pandemic." *Survey Practice* 16 (1). <https://doi.org/10.29115/SP-2023-0011>.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method (Third Edition)*. John Wiley & Sons.

Goel, S., & Salganik, M. J. (2010). Assessing respondent-driven sampling. *Proceedings of the National Academy of Sciences*, 107(15), 6743-6747.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey Methodology (Second Edition)*. John Wiley & Sons.

Groves, R. M., & Heeringa, S. G. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 169(3), 439-457.

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *The Public Opinion Quarterly*, 64(3), 299-308.

# References

---

Heckathorn, D. D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social problems*, 44(2), 174-199.

Peytchev, A., Riley, S., Rosen, J., Murphy, J., & Lindblad, M. (2010, May). Reduction of nonresponse bias through case prioritization. In *Survey Research Methods* (Vol. 4, No. 1, pp. 21-29).

Rao, R. S., Glickman, M. E., & Glynn, R. J. (2008). Stopping rules for surveys with multiple waves of nonrespondent follow-up. *Statistics in Medicine*, 27(12), 2196-2213.

Schouten, B., Peytchev, A., & Wagner, J. (2017). *Adaptive Survey Design*. CRC Press.

Sosenko, F. L., & Bramley, G. (2022). Smartphone-based Respondent Driven Sampling (RDS): A methodological advance in surveying small or 'hard-to-reach' populations. *PLoS ONE*, 17(7), e0270673.

Tourangeau, R., Edwards, B., Johnson, T., Wolter, K., & Bates, N. (Eds.). (2014). *Hard-to-Survey Populations*. Cambridge: Cambridge University Press. doi:10.1017/CB09781139381635.

Wagner, J., & Raghunathan, T. E. (2010). A new stopping rule for surveys. *Statistics in Medicine*, 29(9), 1014-1024.

Wagner, J., West, B. T., Kirgis, N., Lepkowski, J. M., Axinn, W. G., & Ndiaye, S. K. (2012). Use of paradata in a responsive design framework to manage a field data collection. *Journal of Official Statistics*, 28(4), 477.

Wagner, J., Zhang, X., Elliott, M.R., West, B.T., and Coffey, S. (In Press). An Experimental Evaluation of a Stopping Rule Aimed at Maximizing Cost-Quality Tradeoffs. *Journal of the Royal Statistical Society-Series A*.

# References

---

- West, B. T., Chang, W., & Zmich, A. (2021). An experimental evaluation of alternative methods for case prioritization in responsive survey design. *Journal of Survey Statistics and Methodology*, smab031.
- West, B. T., Wagner, J., Hubbard, F., & Gu, H. (2015). The utility of alternative commercial data sources for survey operations and estimation: Evidence from the National Survey of Family Growth. *Journal of Survey Statistics and Methodology*, 3(2), 240-264.
- West, B. T., Zhang, S., Wagner, J., Gatward, R., Saw, H. W., & Axinn, W. G. (2023). Methods for improving participation rates in national self-administered web/mail surveys: Evidence from the United States. *PLoS ONE*, 18(8), e0289695.
- Willis, G. B., Smith, T. W., Shariff-Marco, S., & English, N. (2014). Overview of the special issue on surveying the hard-to-reach. *Journal of Official Statistics*, 30(2), 171-176.
- Zhang, S., West, B.T., Wagner, J., Couper, M.P, Gatward, R., and Axinn, W.G. (2023). Visible cash, a second incentive, and priority mail? An experimental evaluation of mailing strategies for a screening questionnaire in a national push-to-web/mail survey. *Journal of Survey Statistics and Methodology*. DOI: 10.1093/jssam/smac041.