SURVEYS · SAMPLING

River Sampling and Intercept Recruitment on the Web

River sampling recruits survey respondents in real time from the stream of people using websites, apps and social media, rather than from a standing panel. This article explains how river and intercept recruitment work, their strengths, their risks and how to use them well.

ModQu research article · 1,195 words · 13 references (APA 7th edition)

Before the internet, market researchers often stood in shopping centres and stopped passers-by to ask them questions. This **intercept** method was quick and cheap, but it captured only the people who happened to be there, and interviewers chose whom to approach (Sudman, 1980; Bush & Hair, 1985). The web has created a digital version. **River sampling** recruits respondents from the flow of people using websites, apps and social media, inviting them to take a survey at the moment they are online. The term reflects the idea of dipping into a stream of people rather than drawing from a pool of registered panel members (Baker et al., 2013).

How it works

River recruitment usually works through invitations placed on websites, in apps, in online games or in social media feeds. People who click are taken to a short screening questionnaire that checks their eligibility and the quotas still open, and are then routed to a survey that needs respondents like them, often in exchange for a small reward such as points or in-game credit (Callegaro et al., 2014). Figure 1 shows the main steps.

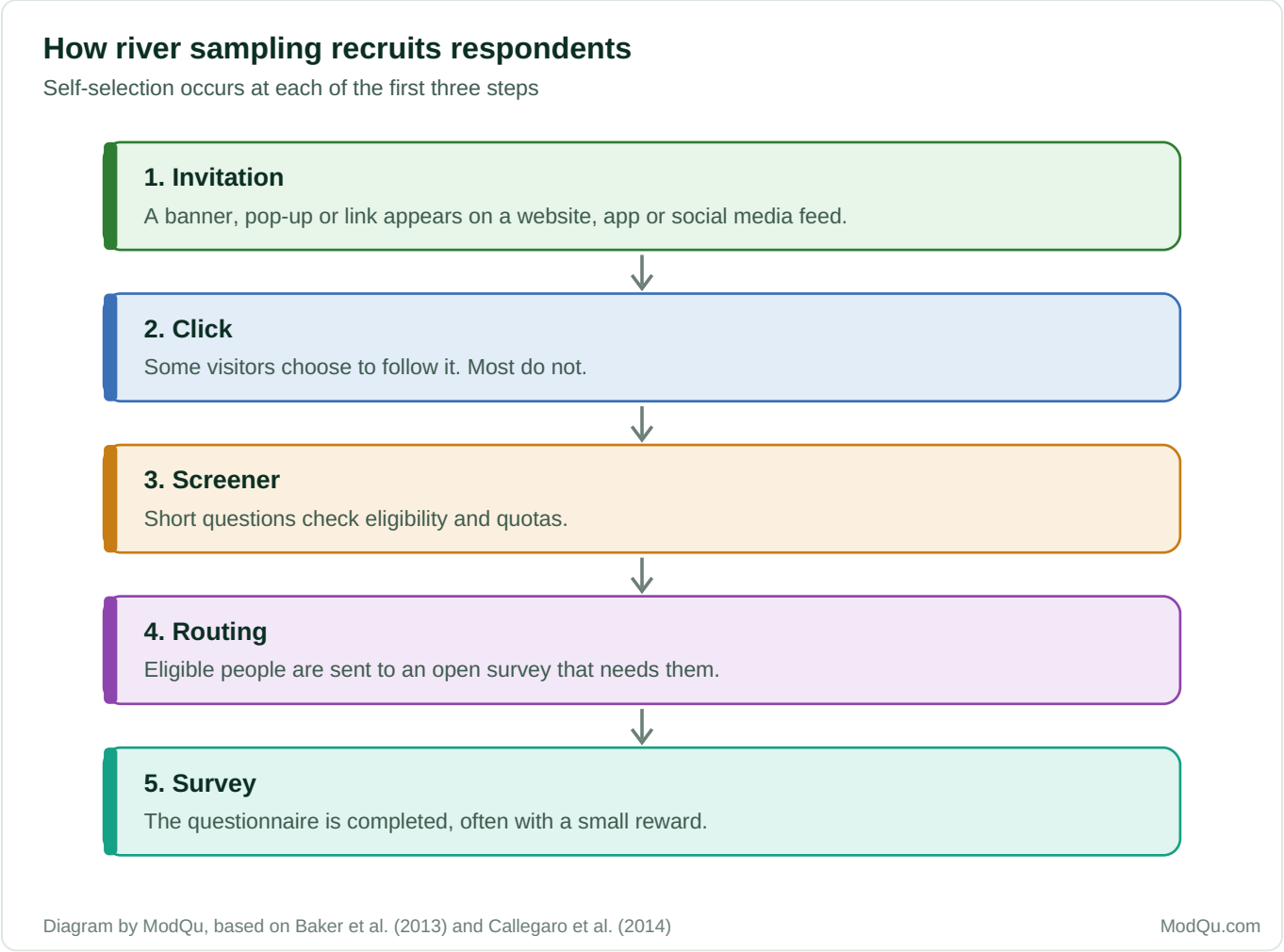


Figure 1. The main steps in river sampling. Diagram by ModQu, based on Baker et al. (2013) and Callegaro et al. (2014).

A related approach is the **website intercept survey**, in which a pop-up invites visitors to a particular website to give feedback, typically about the site itself. Couper (2000) included intercept surveys in his typology of web surveys, noting that they can represent visitors to a site if invitations are issued systematically, for example to every tenth visitor, but not the wider population.

Why it is used

River sampling has several practical advantages. It reaches people who would never join a survey panel, so it can add fresh respondents and reduce dependence on a small group of frequent survey-takers. It can deliver large numbers of respondents quickly. And it can reach people at the moment of an experience, which is valuable for feedback about websites, apps or online purchases (Callegaro et al., 2014).

The risks

The central problem is self-selection. At each step in Figure 1, people decide whether to continue. Those who notice the invitation, click on it and complete the screener differ from those who do not, and their reasons for taking part may be related to the topic of the survey (Bethlehem, 2010). The population reached depends on where the invitations appear: the audience of a gaming site or a news site is not the general population.

Other risks follow from how river traffic is managed. Because respondents are routed to whatever survey needs them, the same stream of people feeds many surveys, and the composition of the sample can change with the time of day, the websites used and the incentives offered. Researchers commissioning a survey often cannot see where their respondents came from (Baker et al., 2013). River samples are also exposed to fraudulent responses and automated programs seeking rewards, which require screening.

Comparisons of online recruitment methods show that the source matters. A study of several online recruitment channels, including search advertising and social media, found that each channel produced samples with different demographic profiles and different costs, and none matched the population closely (Antoun et al., 2016). More generally, studies comparing nonprobability online samples with benchmarks find variable accuracy that depends on the source and the adjustment used (Cornesse et al., 2020).

Place in the wider survey landscape

River sampling grew as response rates to traditional surveys fell and the cost of probability samples rose. Early reviews of internet research already noted that web recruitment offered speed and low cost but at the price of uncertain coverage and self-selection (Fricker & Schonlau, 2002). Survey methodologists have since argued that the future of sampling will involve combining sources of different quality, with probability samples used to calibrate or check cheaper nonprobability sources such as river and panel samples (Brick, 2011). On this view, river samples are most valuable as one input among several, not as a stand-alone substitute for a designed sample.

Using river and intercept samples well

These methods are most defensible when the target population is the audience being intercepted. A survey of a website's visitors, recruited by a systematic pop-up on that website, can represent its visitors reasonably well, apart from nonresponse. Problems grow when river samples are used to represent the general population or a broad market. In that case, good practice includes:

- **Knowing the sources.** Ask which websites and apps supply respondents and in what proportions.
- **Recording the source for each respondent**, so that results can be compared across sources and adjusted.
- **Using quotas and weighting** on variables related to both the topic and the likelihood of taking part (Mercer et al., 2017).
- **Screening for fraud and inattention**, including checks for duplicate respondents and implausible answers.
- **Testing against benchmarks** with known population values.
- **Being transparent** in reports that the sample was recruited through river or intercept methods.

Design considerations

Intercept surveys suffer when invitations are intrusive or appear at the wrong moment. A pop-up that appears before a visitor has completed a task interrupts them and may draw responses mainly from those who are dissatisfied or idle. Timing invitations to appear after a task, keeping them short and making the purpose clear can improve response and reduce bias. Questionnaires should be designed for mobile screens, since many river respondents arrive from phones (Couper, 2008).

Sampling within an intercept survey can also be made more systematic. Rather than showing the invitation to every visitor and accepting whoever responds, a site can show it to a random selection of visitors, for example one in twenty, and record how many were invited. This makes it possible to calculate a response rate and to compare respondents with all invited visitors on information the site already holds, such as pages viewed or device type. Such comparisons show where the respondents differ from the visitors as a whole and allow weighting to correct for it.

An example

An online retailer wants to know how easy its checkout process is to use. It shows a short survey invitation to a random one in twenty customers immediately after they complete a purchase, and records how many were invited, how many responded and the value of their orders. Because the target population is the retailer's own customers who completed checkout, and invitations were systematic, the results represent that population reasonably well. The retailer weights responses to match all invited customers on order value and device type. If it instead wanted to measure the share of all adults who shop online, the same intercept sample would be unsuitable, because it includes only its own customers (illustrative example).

Conclusion

River sampling and web intercepts are the online successors of the shopping-centre intercept. They are fast, cheap and good at reaching people who do not join panels, and they work well when the target population is the audience being intercepted. When used to represent wider populations, they depend on self-selection at several stages, and their accuracy varies with the sources used. Knowing where respondents come from, adjusting for differences and reporting the method honestly are essential (Baker et al., 2013; Lehdonvirta et al., 2021).

References

Antoun, C., Zhang, C., Conrad, F. G., & Schober, M. F. (2016). Comparisons of online recruitment strategies for convenience samples: Craigslist, Google AdWords, Facebook, and Amazon Mechanical Turk. *Field Methods, 28*(3), 231–246. https://doi.org/10.1177/1525822X15603149

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., & Tourangeau, R. (2013). Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology, 1*(2), 90–143. https://doi.org/10.1093/jssam/smt008

Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review, 78*(2), 161–188. https://doi.org/10.1111/j.1751-5823.2010.00112.x

Brick, J. M. (2011). The future of survey sampling. *Public Opinion Quarterly, 75*(5), 872–888. https://doi.org/10.1093/poq/nfr045

Bush, A. J., & Hair, J. F., Jr. (1985). An assessment of the mall intercept as a data collection method. *Journal of Marketing Research, 22*(2), 158–167. https://doi.org/10.1177/002224378502200205

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (Eds.). (2014). *Online panel research: A data quality perspective*. Wiley. https://doi.org/10.1002/9781118763520

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B., & Wenz, A. (2020). A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *Journal of Survey Statistics and Methodology, 8*(1), 4–36. https://doi.org/10.1093/jssam/smz041

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*(4), 464–494. https://doi.org/10.1086/318641

Couper, M. P. (2008). *Designing effective web surveys*. Cambridge University Press. https://doi.org/10.1017/CBO9780511499371

Fricker, R. D., Jr., & Schonlau, M. (2002). Advantages and disadvantages of Internet research surveys: Evidence from the literature. *Field Methods, 14*(4), 347–367. https://doi.org/10.1177/152582202237725

Lehdonvirta, V., Oksanen, A., Räsänen, P., & Blank, G. (2021). Social media, web, and panel surveys: Using non-probability samples in social and policy research. *Policy & Internet, 13*(1), 134–155. https://doi.org/10.1002/poi3.238

Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). Theory and practice in nonprobability surveys: Parallels between causal inference and survey inference. *Public Opinion Quarterly, 81*(S1), 250–271. https://doi.org/10.1093/poq/nfw060

Sudman, S. (1980). Improving the quality of shopping center sampling. *Journal of Marketing Research, 17*(4), 423–431. https://doi.org/10.1177/002224378001700401