

Sampling And Collecting Quantitative And Qualitative Data

Posted on May 29, 2018

Introduction

Sampling and data collection connect a research question to the evidence used to answer it. A technically sophisticated analysis cannot repair a sample that excludes important groups or a measure that does not represent the intended concept. The original discussion correctly distinguished probability and non-probability approaches and identified questionnaires as useful tools. However, probability sampling does not always give every person an equal chance; it gives each eligible unit a *known, nonzero* probability under the selected design. Likewise, validity is broader than generalizability, and a method may be reliable without measuring the intended construct accurately (Teddle & Yu, 2007).

Beginning With the Research Question and Target Population

A study should first define the population about which it seeks to draw conclusions. The target population may be all nurses in a region, adults with a particular diagnosis, students enrolled in a program, or organizations operating in an industry. The accessible population is the portion that can realistically be reached. A sampling frame is the operational list or mechanism used to identify eligible units. Coverage error occurs when the frame omits members, includes ineligible entries, or duplicates people.

The unit of analysis must also be explicit. A study may collect answers from individuals but make claims about households, schools, hospitals, or communities. Confusing the respondent with the analytic unit can produce invalid conclusions. Eligibility criteria, geography, time period, and key definitions should be determined before recruitment rather than adjusted after investigators see the data.

Probability Sampling

Probability sampling uses a random mechanism and known selection probabilities. In simple random sampling, each eligible unit has the same probability of selection. Systematic sampling selects a random starting point and then every k th unit, provided the ordering does not contain a pattern that

biases selection. Stratified sampling divides the population into meaningful groups and samples within each group. Disproportionate stratification can deliberately oversample a smaller group, with survey weights later restoring population estimates.

Cluster sampling selects naturally occurring groups such as schools, villages, clinics, or city blocks. It can reduce travel and listing costs, but people within a cluster often resemble one another, increasing sampling variance. Multistage designs may select regions, then institutions, then individuals. Unequal probabilities, clustering, and nonresponse must be reflected in weights and variance estimation. A probability design supports population inference only when implementation, coverage, and response are also adequate.

Non-Probability Sampling

Non-probability methods are appropriate when the objective is depth, theory development, rare-case access, feasibility testing, or exploration rather than statistical prevalence estimation. Purposive sampling selects information-rich participants according to criteria relevant to the question. Maximum-variation sampling seeks contrasting experiences; criterion sampling requires a shared characteristic; expert sampling recruits specialized knowledge. Snowball or respondent-driven approaches can reach hidden populations through participant networks, although network dependence can bias who is reached.

Convenience sampling is fast but vulnerable to self-selection and coverage bias. Quota sampling can ensure representation on selected characteristics without giving units known probabilities. Qualitative researchers often use theoretical sampling, in which emerging analysis guides the recruitment of cases that can refine or challenge developing concepts. Quality is judged through transparent rationale, sufficiency, reflexivity, and depth rather than pretending the sample is statistically representative.

Sample Size and Information Needs

Quantitative sample size depends on the primary outcome, desired precision, expected variability, effect size, significance threshold, statistical power, design effect, subgroup analyses, and anticipated missing data. A larger sample does not automatically remove bias. Thousands of responses from a badly covered convenience panel may be less credible than a smaller probability sample. Researchers should plan from the analysis and primary endpoint rather than rely on a single rule of thumb.

Qualitative adequacy concerns information power and analytic sufficiency. A focused sample with specific participants, strong interviews, and a narrow question may need fewer cases than a broad, heterogeneous inquiry. “Saturation” should not be used as a magical number. Investigators should state what kind of saturation they sought, how they monitored new information, and why additional recruitment was unlikely to change the interpretation materially.

Quantitative Data Collection

Quantitative methods include questionnaires, structured observations, physiological measurements, administrative records, sensors, and standardized tests. A questionnaire should map each item to a construct and an analytic purpose. Clear wording, balanced response options, appropriate recall periods, logical ordering, and accessible design reduce measurement error. Pilot testing and cognitive interviewing reveal whether respondents interpret items as intended.

Online surveys can be inexpensive and rapid, but they exclude people with limited connectivity, digital literacy, language access, or assistive technology unless alternatives are provided. Anonymity may improve disclosure of sensitive information, but it does not guarantee accuracy. Social desirability, recall error, satisficing, straight-lining, and multiple submissions remain possible. Response rate alone is not a complete quality indicator; differences between respondents and nonrespondents matter more.

Qualitative Data Collection

Interviews, focus groups, participant observation, documents, diaries, photographs, and open-ended digital methods can explore meaning and context. Semi-structured interviews balance consistency with flexibility. Focus groups reveal interaction and shared norms but may suppress dissent or disclosure of sensitive experiences. Observation can compare reported practices with actual routines, while documents show how institutions formally represent decisions.

The researcher is part of the qualitative instrument. Tone, identity, assumptions, language, and relationship to participants influence what is said and how it is interpreted. Reflexive notes, interviewer training, iterative guides, and attention to power differences improve credibility. Recording and verbatim transcription can preserve detail, but consent, data security, and the risks of voice identification require consideration.

Mixed-Methods Sampling and Integration

Mixed-methods research should connect sampling choices to the purpose of integration. In a sequential explanatory design, a quantitative sample may establish a pattern and a purposive qualitative subsample may explain it. In an exploratory design, interviews may identify constructs used to build a survey. Concurrent designs can compare numeric trends with narratives collected during the same period (Onwuegbuzie & Collins, 2007).

The samples may be identical, nested, parallel, or multilevel. Integration should occur in design, analysis, and interpretation—not merely by placing two separate studies in one report. Researchers can use joint displays, case comparisons, or explanatory matrices to show where findings converge,

diverge, or address different dimensions.

Ethical Issues in Recruitment and Collection

Ethical research requires informed and voluntary participation, fair selection, privacy protection, risk minimization, and a meaningful right to stop or skip questions. Incentives should compensate time without becoming coercive. Recruitment through supervisors, teachers, clinicians, or community leaders can create perceived pressure, so roles and consequences must be clarified. Vulnerable populations may require additional protections without being excluded automatically from research relevant to them (American Association for Public Opinion Research, 2023).

Sensitive questions can trigger distress or disclose abuse, illegal activity, health conditions, or immigration status. A protocol should specify warning language, response options, referral resources, and limits of confidentiality. Collecting less identifiable information, separating contact details from responses, encrypting records, and defining deletion schedules reduce privacy risk.

Reliability, Validity, and Trustworthiness

Reliability concerns consistency. Test-retest reliability examines stability over time; inter-rater reliability examines agreement; internal consistency assesses whether items intended to form a scale behave coherently. High reliability does not prove validity. A scale can consistently read the wrong weight, as the original example illustrates (Drost, 2011).

Validity concerns whether evidence and theory support the interpretation and use of scores. Content validity asks whether the measure represents the construct domain. Construct validity examines expected relationships and distinctions. Criterion validity compares with a relevant external outcome. Internal validity concerns causal inference, while external validity concerns transfer or generalization across populations and settings. Generalizability is therefore only one dimension.

Qualitative researchers often discuss credibility, dependability, confirmability, and transferability. Strategies include prolonged engagement, triangulation, negative-case analysis, audit trails, peer debriefing, and rich contextual description. Member reflection can be useful, but participants do not hold sole authority over an interpretation involving broader evidence.

Missing Data, Bias, and Transparent Reporting

Missingness can result from skipped items, loss to follow-up, device failure, or linkage problems. Deleting incomplete cases can bias estimates when missingness relates to the outcome or exposure. Prevention through concise instruments, reminders, flexible modes, and careful tracking is preferable. Analysis may require weighting, multiple imputation, sensitivity analysis, or explicit qualitative

treatment of absent voices.

A strong report identifies the frame, recruitment dates, sampling design, inclusion criteria, response and attrition, data-collection mode, instrument development, weighting, interviewer procedures, and deviations from protocol. Limitations should describe the direction and likely consequence of bias rather than merely stating that “the sample was small.”

Mode Effects and Inclusive Administration

The same question can produce different answers when asked face to face, by telephone, on paper, or online. Interviewer presence may increase social-desirability pressure, while self-administration may improve privacy but reduce opportunities for clarification. Mixed-mode studies should test whether wording, visual layout, and response options remain comparable. Accessibility requires screen-reader compatibility, adequate contrast, keyboard navigation, plain language, translated materials, and alternatives for people who cannot use the dominant mode.

Fieldwork monitoring should examine response patterns by subgroup, interviewer, site, and time. Unusually short completion times, duplicate records, excessive missingness, or identical response strings may indicate poor-quality data, but automatic deletion rules can also remove legitimate respondents. Quality control should be specified before investigators know which records support their hypothesis.

Preregistration and Separation of Confirmatory From Exploratory Work

Sampling decisions can be manipulated unintentionally when researchers change eligibility, outcomes, or subgroup definitions after seeing results. Preregistration records the planned population, sample size, exclusions, measures, and analysis before outcome inspection. It does not prevent discovery; it distinguishes confirmatory tests from exploratory findings that require replication.

Qualitative inquiry is often iterative, but transparency remains possible. Researchers can document when the interview guide changed, why new cases were sought, and how emerging interpretations affected sampling. An audit trail makes adaptation visible rather than pretending the final design existed from the beginning.

Conclusion

Sampling is not a preliminary administrative step; it defines whose evidence can shape the

conclusion. Probability methods support population estimates when selection probabilities, coverage, implementation, and weighting are sound. Non-probability methods can produce valuable explanatory insight when their purpose and limits are transparent. Questionnaires, interviews, observations, and records each create distinct opportunities and errors. Ethical recruitment, accessible collection, valid measurement, reflexive analysis, and honest reporting are necessary for both quantitative and qualitative credibility (Teddle & Yu, 2007).

References

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives*, 38(1), 105–124.

Onwuegbuzie, A. J., & Collins, K. M. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, 12(2), 281–316.

Teddle, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, 1(1), 77–100.

American Association for Public Opinion Research. (2023). *Standard definitions: Final dispositions of case codes and outcome rates for surveys*.