

Assessing Research Protocols: Primary Data Collection

By: Maude Laberge, PhD

Definition

Data collection refers to the process in which researchers prepare and collect data required.

The data can be gathered from various sources. The data collected can be transformed into information, which can support decision making. Primary data is original data that is collected for a specific research project.

Advantages:

- Relevance of data to answer the research question: because the data is collected specifically for a research project, there is an alignment between the data collected and the questions that need to be answered.
- Researchers are familiar with the data: the researchers define in detail what is going to be collected and the study population and may participate (or at least supervise) the data collection process which means that they are knowledgeable of what is being collected, how and from whom and the potential issues in the data.
- Timely: collected in the context of the research project. This means that the data can be available immediately after it is collected. In contrast, secondary use of data may require permission and may only become available after an extended period of time after the data was collected.
- Addresses the specific objectives of the research: data collection the method, the sample population it is collected from and the data elements collected correspond to the identified objectives of the study.

Disadvantages:

- Expensive: data that is being collected specifically for the purpose of a study may be quite expensive to collect. Depending on the methodology, it may require the development of a tool and its distribution by mail; financial incentives may be included to increase response rates; it may require resources to call survey respondents. The tool needs to be in a format that allows for ease of the collection but may require transformation for the analysis.
- Time consuming: identifying the population and recruiting participants may take a long time; low response rates means that the period for recruitment may be extended. Reminders may need to be sent to participants to increase response.
- Challenging: there are many challenges in primary data collection that are specific to the context in terms of the data that is being collected and from whom. After identifying

the population of interest, one needs to consider the availability of the respondents/participants (executives may not have much time to participate) and the recruitment strategy (to identify who should be participating in the study given its objectives). There may also be important ethical issues with vulnerable groups (for instance if the study population consists of people who are homeless, etc).

How will it be collected?

Methods to collect primary data: surveys (personal interview, mail, in-house, self-administered, telephone, fax, email, web), focus groups, individual in-depth interviews, experiments, observation, simulation.

When will it be collected & over which time period?

Cross-sectional data: one specific period of time

Time-series: the same data is collected from a different group of people over defined period of time, usually with the same interval between periods of data collection; the respondents are different from one collection period to the next but should represent the same population

Panel: multiple data collection periods where the same data is collected from the same sample group. This allows to see changes within that group and analyze the potential effect of an intervention on the group.

From whom? Define target participants

Identify the target population: what the characteristics of the target population are including the size and its availability for the study; define why these characteristics are important in relation to the study objectives. Determine if there are published lists of potential respondents and how to get access to the population of interest.

From how many people? Define sample

Sampling methodology: define whether it will be a single or multistage sampling. “Identify whether the sampling design for this population is single or multistage (called clustering). Cluster sampling is ideal when it is impossible or impractical to compile a list of the elements composing the population. A single-stage sampling procedure is one in which the researcher has access to names in the population and can sample the people (or other elements) directly. In a multistage, clustering procedure, the researcher first samples groups or organizations (or clusters), obtains names of individuals within groups or clusters, and then samples within the clusters” (Creswell 2003).

How? Recruitment strategy and selection process

Selection process for individuals. The sample can be random, i.e. each individual in the population has an equal probability of being selected (a systematic or probabilistic sample), or it can be a convenience sample, i.e. respondents are chosen based on their convenience and availability. With randomization, a representative sample from a population provides the ability to generalize to a population.

Stratification: Stratification means that specific characteristics of individuals (e.g., both females and males) are represented in the sample and the sample reflects the true proportion of individuals with certain characteristics of the population. When randomly selecting people from a population, these characteristics may or may not be present in the sample in the same proportions as in the population; stratification ensures their representation. Also identify the characteristics used in stratifying the population (e.g., gender, income levels, education). Within each stratum, identify whether the sample contains individuals with the characteristic in the same proportion as the characteristic appears in the entire population.

Data Collection Procedures in Qualitative Studies

Where will data be collected from?

Identification of selected sites or individuals. The idea behind qualitative research is to purposefully select participants or sites (or documents or visual material) that will best help the researcher understand the problem and the research question. This does not necessarily suggest random sampling or selection of a large number of participants and sites, as typically found in quantitative research. A discussion about participants and site might include four aspects: the setting (where the research will take place), the actors (who will be observed or interviewed), the events (what the actors will be observed or interviewed doing), and the process (the evolving nature of events undertaken by the actors within the setting).

What type of data will be collected?

1. Field notes from observations: the researcher takes field notes on the behavior and activities of individuals at the research site. In these field notes, the researcher records, in an unstructured or semi-structured (using some prior questions that the inquirer wants to know) way, activities at the research site. The qualitative observer may also engage in roles varying from a non participant to a complete participant.
2. Views and opinions from interviews: the researcher conducts face-to-face interviews with participants, interviews participants by telephone, or engages in focus group interviews with six to eight interviewees in each group. These interviews involve unstructured and generally open-ended questions that are few in number and intended to elicit views and opinions from the participants.
3. Documents: During the process of research the qualitative investigator may collect documents. These may be public documents (e.g., newspapers, minutes of meetings, official reports or private documents (e.g., personal journals and diaries, letters, e-mails).

4. Audio and visual material: this data may take the form of photographs, art objects, videotapes, or any forms of sound.

Summary Checklist:

- ✓ Define data to be collected
- ✓ Determine need/use of the data to be collected
- ✓ Determine if sampling is necessary and if so, sampling technique
- ✓ Identify data sources
- ✓ Determine methods of data collection
- ✓ Consent from participants
- ✓ Develop a data collection tool (survey, interview guide, etc.)
- ✓ Test & validate the tool
- ✓ Determine how the data will be recorded
- ✓ Document data collection method

Table 1. A list of Qualitative Data collection Approaches

- Gather observational notes by conducting on observation as a participant
- Gather observational notes by conducting on observation as an observer
- Conduct an unstructured, open-ended interview and take interview notes
- Conduct an unstructured, open-ended interview, audiotape the interview, and transcribe the interview
- Keep a journal during the research study
- Have a participant keep a journal during the research study
- Optically scan newspaper accounts
- Collect personal letters from participants
- Analyze public documents (e.g., official memos, minutes, records, archival material)
- Examine autobiographies and biographies
- Have a participant write her or his autobiography
- Write your own (the researcher's) autobiography
- Have participants take photographs or videotapes (i.e., photo elicitation)
- Examine physical trace evidence (e.g., footprints in the snow)
- Videotape a social situation or an individual/group
- Examine photographs or videotapes
- Collect sounds (e.g., musical sounds, a child's laughter, car horns honking)
- Collect e-mail or electronic messages
- Examine possessions or ritual objects to elicit views during an interview
- Collect smells, tastes, or sensations through touch

References

Cochrane Effective Practice and Organisation of Care Review Group (EPOC). Data Collection Checklist

<https://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/datacollectionchecklist.pdf>, consulted October 27 2014

Creswell JW. (2003) Research design: qualitative, quantitative, and mixed methods approaches (2nd Edition), Thousand Oaks, CA: Sage Publications.

Hox JJ, Boeije HR. (2005) Data collection, Primary vs. Secondary. Encyclopedia of Social Measurement, Volume 1:593-9

Kothari CR. (2011) Research Methodology: Methods and Techniques (2nd Edition), New Age International

Polonsky MJ, Waller DS. (2010) Designing and Managing a Research project: A Business Student's Guide, Sage Publications

World Bank Group. Primary Data Collection Methods.

<https://www.wbginvestmentclimate.org/toolkits/public-policy-toolkit/upload/Primary-Data-Collection-Techniques-for-Toolkit.pdf> , consulted October 27th, 2014