Survey Research

Abstract

Survey research is a non-experimental research approach used to gather information about the incidence and distribution of, and the relationships that exist between, variables in a pre-determined population. Its uses include the gathering of data related to attitudes, behaviours and the incidence of events. Survey research in one form or another has existed for over two millennia with the population census of Caesar Augustus (St. Luke's Gospel) being an early example. For most modern researchers sample surveys are more cost effective and easier to undertake than population surveys when gathering information; however, this increases the risk of both representation and measurement errors. There are a number of different forms of survey research; however they all share common steps and common limitations. The purpose of this article is to discuss these steps with a view to highlighting some of the common difficulties.

Survey Research

Quantitative research is composed of two distinct yet methodologically interconnected research approaches; experimental and survey research (Davis, 2007). Survey research is a non-experimental research approach used to gather information about the incidence and distribution of, and the relationships that exist between, variables in a pre-determined population. Surveys, depending on the data required, use both total populations (population census) and samples of the population to gather this data. This approach to gathering data pre-dates modern research and has been documented throughout history where it has been used to gather information about people and wealth. Some of the earliest references to surveys include Caesar Augustus (b. 63BC d. 27AD) undertaking a population census to determine taxation (St. Luke's Gospel), and the 'Domesday Book' which was completed for William the Conqueror in 1086 and was a survey of the wealth of his kingdom again to determine taxes. In the modern era, population censuses are still used and a familiar example is the decennial census, which modern states use to gather data such as the size and age of their populations. However, for most researchers, population surveys are too costly and too difficult to undertake, so instead they use sample surveys to gather data from a group of participants with the view to generalising this information to the population in question.

Survey research can take a number of forms. Descriptive survey research is the simplest form of survey. Its uses include gathering data related to attitudes, behaviours and the incidence of events. This type of survey offers a snapshot of the phenomenon being studied and is uncomplicated as it usually only involves single contact with the sample being studied. However because it is a snapshot, it does not allow for changes that may occur due to unforeseen variables (McKenna et al, 2006). Longitudinal studies, where the survey is administered a number of times over the period of the research, can be useful in overcoming this problem. Correlational and comparative surveys are used to study and compare the relationships between variables. However, whichever type of survey is used, they all share common steps and common limitations. The purpose of this article is to discuss these steps with a view to highlighting some of the common difficulties.

Defining the Problem / Formulating a Research Question

The first stage of survey research is the identification of an explicit research problem. In survey research, as in all quantitative research, it is important that the researcher is clear as to the purpose of the study and has a precise research question. A research study that seeks to answer numerous diverse questions rather than an explicit question is often weak. This is the case particularly where the researcher simply asks questions because they seem interesting and then ends up dredging through the results looking for useful findings (Umbach, 2005; Kelly et al, 2003). Hallberg (2008) suggests that when researchers identify a research problem, they should then attempt to determine which factors influence this phenomenon. He firstly recommends a thorough review of the literature to identify factors that can affect the problem. A theoretical framework of influencing variables can then be constructed to help them decide which variables need to be controlled and which must be included to achieve findings that are likely to be valid and reliable. While it may seem that the inclusion of as many variables as possible may create a more realistic portrait of the phenomenon being studied, surveys that are too long have both a negative effect on response rates and can be unwieldy to manage (Hallberg, 2008). It is thus equally as important to determine which

variables can be excluded as it is to decide which must be included.

Data Collection Methods

There are a variety of data collection methods associated with survey research, the most common being self-administered questionnaires and structured interviews. Non-participant observation may also be used but is employed infrequently (McKenna et al. 2006). A relatively new and popular method of data gathering is through computer technology with some suggestions that this may supersede traditional data gathering methods in the future (Truell, 2003).

Self-Administered Questionnaires

Traditionally self-administered questionnaires were the backbone of survey research in that they allowed, especially in the case of postal questionnaires, large numbers of individuals from widespread geographical locations to be sampled cost effectively (Polit and Beck, 2008). However, the major limitation to this approach has always been poor response rates, which can restrict researchers in their quest to generalize findings to the population. A 50% response rate is generally regarded as acceptable, 60% is regarded as good and a 70% response rate is usually regarded as very good (Babbie, 1990), however, response rates are often as low as 30% (Gerrish and Guillaume, 2006). It is also important to note that in an era where the demands for survey generated data and generalizable data is increasing the response rates, in both the United States and Europe, are continuing to fall (Porter, 2004a). Attempting generalizations in the face of low response rates can lead to what Umbach (2005) identifies as non-response errors. He states that these result from individual or item non-responses. Individual non-response errors occur when insufficient numbers of the sample respond, and a probable non-representativeness of the population results. Item non-response errors occur when respondents do not answer one or more questions (Umbach, 2005).

Researchers' attempts to overcome the problem of low response fall into two categories. The first involves the use of techniques to persuade more sample members to respond (Table 1) and the second entails weighting the survey for non-responses (Porter, 2004a). This latter process increases the complexity of the approach and despite this may still not be sufficient to prevent rejection of the findings of a study if the response rate is too low (Porter 2004a).

Other limitations that are associated with self-administered questionnaires are the possibility that the respondent either did not complete the questionnaire himself or sought help to do so. This may interfere with the representativeness of the sample particularly if it happened frequently within a study. The difficulty is that it is not possible to know if it happens. Difficulty reading, interpreting words or writing can exclude particular groups from participation and thus may bias the results (Hallberg, 2008).

Structured Interviews

A structured interview is a process whereby a researcher administers a questionnaire, known as an interview schedule, to the respondent directly. The researcher can interview the respondent face to face, or via a medium such as the telephone or the internet (Parahoo, 2006). An interview schedule, like a questionnaire, seeks answers to questions in a precise, preordained order and the interviewer follows the predetermined sequence (McKenna et al, 2006). Structured interviews have some advantages over self-administered questionnaires.

Interviews generally achieve better response rates. This may be due to reluctance on the part of the potential respondent, when approached, to refuse the interviewer (Babbie, 1990). It may also be due to an increased opportunity for the interviewer to promote the purpose of the research (Kelly et al, 2003). This data gathering method can also be useful in reducing item non-responses and the number of "Don't Know's" as the interviewer is in a position to clarify any questions that are unclear to the respondent. However, in clarifying items the interviewer must ensure that the meaning of the question is not altered (Parahoo, 2006). Structured interviews are also of benefit in gathering data from groups who have difficulties with reading and/or writing, or respondents who are too young or infirm to complete self-report questionnaires. In addition, they have the added benefit of preventing the respondent seeking assistance in answering or reading ahead in an attempt to anticipate the 'correct' response (Parahoo, 2006).

Conversely, if undertaken face to face, structured interviews can be costly and time consuming (Kelly et al, 2003). Telephone interviews, on the other hand, offer many of the benefits of face to face interviews at a reduced cost; however, the number of refusals to participate are higher than in the case of face to face interviews but still less than in self administered questionnaires (Kelly et al. 2003). Nevertheless, where the problem being investigated is one that is of a personal or sensitive nature self-administered questionnaires are still the most appropriate as they offer the opportunity for the respondent to remain anonymous (Parahoo, 2006).

Online Data Gathering

The use of computers and the internet for data gathering is still relatively new within the field of research. However, with more individuals becoming computer literate and gaining broadband access to the internet it is progressively becoming the more popular method of choice (Granello and Wheaton, 2004). There are a number of ways that data can be gathered using computers including questionnaires administered via e-mail or via the World Wide Web (Stewart, 2003), Computer Assisted Personal Interviews (CAPI), Computer Assisted Telephone Interviews (CATI), and Audio-Computer Assisted Self-Interview (Audio-CASI) (Polit and Beck, 2008).

The simplest method of employing computers in survey research is sending questionnaires by e-mail. The questionnaire can be sent as an attachment or as part of an e-mail, to be completed and returned. Alternatively, the email can be used to direct the potential respondent to a website where the questionnaire can be accessed (Porter, 2004b). The biggest advantage of electronic questionnaires is cheap access to a large sample and ease of data management (Granello and Wheaton, 2004; Umbach, 2004). Electronic questionnaires incur no costs for stationary, printing or postage although some initial cost may be incurred setting up a web site. Data received can be loaded directly to spread sheets reducing both time and the risk of error. Another advantage is faster turnaround time (Umbach, 2004). Traditional questionnaires can take days to arrive at their destination and the same return time. With electronic questionnaires, arrival is practically instantaneous and on completion, the return time is similar.

Poor response rate appears to be the major disadvantage associated with these approaches to data gathering (Granello and Wheaton, 2004). In the case of e-mailed questionnaires, this may be due to the obvious lack of anonymity as the email records the senders details (Stewart, 2003). Alternatively, Truell (2003) would claim that the response rates vary between studies. Nonetheless, what does seem apparent is multiple follow-ups are required

to reduce the risk of non-response errors. Another disadvantage is the potential for sample bias. Access to the internet is not equally distributed throughout the population, for example in the lower socio-economic groups, and where it is available not everyone is computer literate, e.g. some older individuals; thus, some groups are liable to non-inclusion in the study. There are indications that computer access and literacy is changing; however, the onus is still on researchers to be familiar with the defined population and to ensure that access to, and knowledge of the required technology is available to all within the selected sample (Granello and Wheaton, 2004). Other difficulties identified relate to privacy and security. There is a view that 'spamming' - sending unsolicited e-mails, which includes mass e-mailings - is an invasion of privacy (Umbach, 2004). This has become such a problem that many institutions now employ e-mail filters to block this 'spam mail'. In relation to protection of data, and respondents' identities and personal information, despite the best precautions, absolute security is impossible to guarantee.

CAPI is similar to traditional face-to-face interviews, but the conventional pen and paper is replaced by computer (Lumbantobing et al. 2001). The interviewer usually inputs the participant's response in a coded format that can be uploaded directly to the main data file when the interview is complete (Polit and Beck, 2008). CATI is one of the oldest methods of computer assisted interviewing and similar to CAPI but the interview occurs via the telephone (Lumbantobing et al. 2001). The majority of telephone surveys now use this method of interview (Polit and Beck, 2008). Audio-CASI is a method of self-reporting in which the participant uses a computer to respond to questions recorded by an interviewer. The responses are usually coded numerically, which aids both the respondent and researcher. The method is particularly useful for sensitive topics, as the respondent does not meet the interviewer and thus has a sense of anonymity, and/or where the individual has difficulty reading or writing (Polit and Beck, 2008).

Observation

Observational surveys involve counting discernible objects or incidents identified on an observation checklist or rating scale (Sapsford, 2007). Checklists are a list of behaviours or items that are to be observed, with a space for noting the frequency of the event. Rating scales can be used either in association with checklists or independently and involve evaluating the event in some way (Polit and Beck, 2008); for example an observer may use a checklist to record how frequently a client had an aggressive outburst, and use a rating scale to note the degree of aggression observed.

Similar to other data gathering methods, observation also has limitations. The difficulties here can arise from the participant changing behaviour in an attempt to look good to the observer; this is termed a 'Social Desirability Bias'. Observer biases include enhancement of a contrast effect and central tendency, where the observers adjust ratings either towards a distinctive behaviour or towards the midpoint. Careful construction of checklists and rating scales and good pre-observation training can help to reduce some of these biases (Polit and Beck, 2008).

Considering an Instrument for Data Collection

Whatever the approach to data gathering, it is imperative that the items on the research instrument are developed carefully. At this stage, researchers have the choice of selecting an instrument that has been previously designed and tested or constructing a new instrument for

their study (Hallberg, 2008). Developing a new instrument that is valid and reliable can be both a time consuming and costly undertaking. It is worthwhile undertaking an in-depth review of the literature to identify if a suitable tool is available that has been psychometrically tested (Kelly et al. 2003). Psychometric testing of research instruments involves ensuring that the instrument is both valid, that it measures what it is designed to measure, and reliable, that it is consistent in that measurement. Failure to achieve validity and reliability can lead to measurement errors where the participant responses do not relate to the research question, are open to misinterpretation or there is no homogeneity when compared to other participants responses (Umbach, 2005).

To reduce the risk of measurement errors it is important, as stated previously, to identify clearly the research problem to be investigated and ensure the instrument measures this. It is then necessary to ensure that questions are unambiguous and that they measure all the appropriate attributes of the phenomenon being studied. Ambiguity arises where questions contain double negatives, or are 'double barrelled'. Double barrel questions contain more than one question in the stem and are usually associated with the use of 'and' or 'or' in the item (Davis, 2007). The ambiguity for the respondent is determining which question to answer.

Before administering a data gathering instrument it is important to evaluate its internal validity and reliability. This is particularly the case where it is a newly developed questionnaire or where a previously tested questionnaire is to be used with a different cultural group or environment. Umbach (2005) recommends submitting the questionnaire to both subject area experts and to experts on survey design to test both content and design validity. Reliability tests include 'Test-Retest', which measures the stability of the instrument, and 'Cronbach's Alpha', which is used to evaluate the internal consistency or homogeneity of items that measure the same attribute (Polit and Beck, 2008). Pilot studies can offer useful insights into difficulties encountered by respondents and their views on the format and language of an instrument if the researcher uses the opportunity to gather this data (Parahoo, 2006).

As well as internal validity, external validity also needs to be considered. External validity is a measure of the degree to which the findings of a study can be generalized to similar populations or settings (Polit and Beck, 2008). Threats to this type of validity are related to how participants are selected to participate in the survey.

Sample Selection

History indicates that poor sample selection can lead to findings that are not generalizable. One example is the 'Literary Digest Poll' of 1936 which using telephone directories and car registrations to develop a sample frame, and polling two million voters, predicted that Alfred M. Landon would win the American Presidential race over Franklin D. Roosevelt. Gallup, pioneering the use of quota sampling correctly predicted that Roosevelt would win. However, twelve years later still using quota sampling, Gallup wrongly predicted that Thomas E. Dewey would defeat Harry S. Truman in his attempt to achieve a second term in office. At that time, concern had begun to arise regarding the ability of quota sampling to accurately predict outcomes, and a new approach called probability sampling, which some academic researchers were experimenting with, was used to correctly predict the outcome (Babbie, 1990).

The main difficulties that arose with the 1936 Literary Digest Poll, and the 1948 Gallup Poll were related to the non-representativeness of the samples thus leading to poor external

validity. The use of probability sampling increases the odds that the sample will be representative of the population and the likelihood that the findings can be generalized. However, like a well-known lager that claims that, it is 'probably the best in the world'; probability sampling only claims to be probably representative – representativeness is not guaranteed. Probability sampling is an approach to sample selection, which ensures that all units (e.g. individuals) in the population have an equal opportunity, which is greater than zero, of being selected to participate. There are a number of techniques known as random sampling used to select a probability sample (Burns and Grove, 2007).

The most commonly used type of random sampling is 'Simple Random Sampling'. Depending on the complexity of the sample, a simple random sample can be achieved as easily as putting all the names in the sample frame into a hat and drawing out the required number. For larger samples and sample frames, samples can be selected by giving a number to each member of the sample frame and selecting the numbered unit using a table of random numbers or instructing a computer to randomly select the units (Burns and Grove, 2007).

When the sample required is more complex, for example where it is necessary to ensure representativeness of interdisciplinary groups in the sample frame (Occupational, Speech and Language and Physiotherapists) a 'Stratified Random Sample' is used. This method of sampling needs a more detailed sampling frame, as it has to identify the individuals in each group, and what proportion of the population they represent. Once this information has been acquired, the sample is randomly selected ensuring that each group is proportionately represented (Parahoo, 2007).

'Systematic Random Sampling' is a method that employs 'systematic intervals' to select individuals from a sample frame. If a sample of 400 were needed from a sample frame of 4,000 then the interval would be every 10^{th} individual. To ensure randomness the first individual should be selected using a table of random numbers (Polit and Beck, 2008) and ideally, the sample frame should be random in its organisation.

'Cluster sampling', which is also known as 'multi-stage sampling', is a method used to randomly sample from large geographical or national areas. The starting point however is not individuals, but sites from which samples can be derived. If a national survey of nurses working in hospital were to be undertaken, the researcher, using a sample frame of all hospitals, would randomly select a number of these sites. Once these sites had been identified, then individual nurses from those hospitals would be randomly selected to participate in the study (Davis, 2007). As all hospitals were initially included, then all nurses who worked in those hospitals had an equal chance greater than zero of being selected to participate, and thus the sampling is probably representative.

Non-Probability samples such as convenience or quota samples are much easier to obtain than random samples; however, from a statistical perspective, they are regarded as having a low probability of population representativeness. On the other hand, even probability samples are at risk of sampling error; though the researcher can reduce this risk by having an adequate sample size. An adequate sample for statistical analysis can be determined using power analysis. To be acceptable the level of power must be at least 0.8 which offers an 80% probability of overcoming Type II errors (failing to recognise relationships that actually exist) (Burns and Grove, 2007). Generally when using probability sampling the larger the sample the lesser the degree of sampling error (Polit and Beck, 2008).

Ethical Considerations

Researchers have an obligation to uphold the ethical rights of the participants in their study. Two important ethical considerations relate to informed consent and confidentiality. Before asking subjects to participate in a research study, the researcher should inform them regarding the purpose of the study, their right to refuse to participate and, in the case of interviews or observational data gathering, withdraw from the study. Potential participants should also be informed whether their identities are anonymous or how the researcher will ensure confidentially. Participants are usually asked for written consent in the case of face-to-face interviews or observations, whereas in the case of questionnaires consent is usually regarded as implicit in the completion and return of the questionnaire (Hallberg, 2008). Ethics committees may include other stipulations in relation to upholding participants' rights depending on the purpose of the research and how the data is being gathered.

Data Analysis

The purpose of data analysis is to present the vast quantity of gathered data in a summarised way that is comprehensible to the reader. In survey research data are presented through statistics. There are two types of statistics: Descriptive and Inferential. Descriptive statistics use frequency distributions, central tendency and variability to describe single variables. Bivariate descriptive statistics, such as contingency tables and correlations, can be used to show two variable relationships (Polit and Beck, 2008). Inferential statistics use either parametric (e.g. *t*-test independent or paired, analysis of variance, repeated measures, Pearson's product-moment correlation) or non-parametric (e.g. Mann-Whitney U-test, Chi-square test, Kruskal-Wallis test) tests to seek relationships between variables with a view to predicting outcomes (Parahoo, 2006). It is important to mention here some potential errors that can occur. Adjustment error can occur due to item non-responses and the managing of missing data (Umbach, 2005). Researchers need to identify what, if any, data is missing and determine whether they will manage this through traditional strategies such as exclusion or deletion, or use statistical algorithms (Croninger and Douglas, 2005). Another consideration is processing error, which can occur due to miscoding open-ended questions, data entry errors and outliers (Umbach, 2005). Additionally, researchers should recognize that trawling blindly through data looking for significant relationships can lead to false positives (Type I errors).. This is especially the case when large numbers of variables and outcomes are involved and significance is only set at $P<0.05$ (Smith and Ebrahim, 2002). Careful planning can help researchers avoid some of the problems associated with data analysis and reduce the risk of potential errors occurring.

Conclusion

Survey research in one form or another has existed for over two millennia. In the modern era with greater population sizes and wider territorial distribution of individuals, sampling has generally replaced census surveys due to cost. This same financial constraint also applies to time spent on formulating research questions, planning data collection, developing research tools and data analysis. Newer survey methods of data collection may also help in reducing cost but still tend to have high non-response rates. Thus in reducing cost there is the likelihood of a greater risk of errors occurring in either representation or measurement. Researchers need to be aware of these errors and reduce the risk in those areas that they can control.

Key Words

Survey Research / Postal Questionnaires / Measurement and Representation Errors / Sampling.

References

Babbie E (1990) *Survey Research Methods* 2nd Ed. Wadsworth Publishing Co., Belmont Ca.

Burns N, Grove S (2007) *Understanding Nursing Research: Building and Evidence Based Practice* 4th Ed. Saunders, St. Louis

Croninger R G, Douglas K M (2005) Missing Data and Institutional Research. *New Directions for Institutional Research* **2005**(127): 33-49.

Davis M B (2007) *Doing a Successful Research Project.* Palgrave, Houndmills.

Gerrish K, Guillaume L (2006) Whither survey research? The challenge of undertaking postal surveys within the UK research governance framework. *Journal of Research in Nursing* **11**(6): 485-497.

Granello D H, Wheaton J E (2004) Online data collection: Strategies for Research. *Journal of Counseling and Development* **82**(4): 387-393.

Hallberg I R (2008) Surveys. In Watson R, McKenna H, Cowman S, Keady J, eds. *Nursing Research: Design and Methods* Elsevier, Edinburgh: 179-187.

Kelley K, Clark B, Brown V, Sitzia J (2003) Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care* **15**(3): 261-266.

Lumbantobing S, Nachrowi N, Harforo W (2001) *Guidelines on the Application of New Technology to Population Data Collection and Capture.* http://www.unescap.org/stat/pop-it/pop-guide/pop-guide.asp#guide-capture

McKenna H, Hasson F, Keeney S (2006) Surveys. In Gerrish K, Lacey A eds. *The Research Process in Nursing* 5th Ed. Blackwell Publishing, Oxford: 260-273.

Parahoo K (2006) *Nursing Research: Principles, Process and Issues* 2nd Ed. Palgrave Macmillan, Houndmills.

Polit D F, Beck C T (2008) *Nursing Research: Generating and Assessing Evidence for Nursing Practice* 8th Ed. Wolters Kluwer | Lippincott Williams & Wilkins, Philadelphia.

Porter S R (2004a) Raising response rates: What works? *New Directions for Institutional Research* **2004** (121): 5-21.

Porter S R (2004b) Pros and cons of paper and electronic surveys. *New Directions for Institutional Research* **2004**(121): 91-97.

Sapsford R (2007) *Survey Research* 2nd Ed. Sage Publications, London.

Smith G D, Ebrahim S (2002) Data dredging, bias or confounding. *British Medical Journal* **325**(7378): 1437-1438.

Stewart S (2003) Casting the net: Using the internet for survey research. *British Journal of Midwifery* **11**(9): 543-546.

Truell A D (2003) Use of internet tools for survey research. *Information Technology, Learning and Performance Journal* **21**(1): 31-37.

Umbach P D (2004) Web surveys: Best Practice. *New Directions for Institutional Research* **2004**(121): 23-38.

Umbach PD (2005) Getting back to the basics of survey research. *New Directions for Institutional Research* **2005**(127): 91-100.

Table 1: Increasing Response Rates

Problem	Suggested Solutions	Rationale
Individual Non-Response	**Multiple contacts with subjects** including pre-notification of an impending survey questionnaire, the questionnaire, follow up post card, and second questionnaire.	Pre-notification decreases the chance that the subject will discard the questionnaire in error. The questionnaire and stamped addressed return envelope should follow within one to two weeks and contain a personalised cover letter. Post card follow-ups are useful for those who have forgotten to complete the questionnaire. These are usually sent about two weeks following distribution of the questionnaire. If following this response rates are still low a second questionnaire, cover letter and stamped addressed envelope can be sent. Sending multiple questionnaires does improve response rates, but can be costly.
	Personalising contacts	When individuals feel that their personal assistance has been requested, they are more likely to want to participate. Similarly if they feel that they have been selected to be a part of an exclusive group to participate in a study, this uniqueness is more likely to lead to participation.
	Advance Monetary incentives / Rewards	Offering small rewards or small monetary incentives such as a pen or euro/pound invokes a reciprocal response from the respondent who feels they are returning the favour. However, the reward should be of a small enough value not to be regarded as a bribe or compensation. Promises of rewards when the questionnaire is completed and returned have little effect on the response rate.
Item Non-Response	**Short questionnaire**	Shorter questionnaires tend to obtain better overall response rates and less item non-responses.
	Clear concise items	Items that are unclear or imprecise can confuse the respondent who then skips that item.

(Porter, 2004a; Umbach, 2005)