Strengths and weaknesses of online surveys

M. Siva Durga Prasad Nayak¹, K.A. Narayan²

¹(Tutor, Department of Community Medicine, Government Medical College, Ongole / NTR University of Health Sciences, India)
²(Professor, Department of Community Medicine, Mahatma Gandhi Medical College and Research Institute / Sri Balaji Vidyapeeth, India)

Abstract: Data is of paramount importance for research. Though the methodology for research studies may vary, every research is based on data that should be of good quality and then analysed and interpreted to yield information. The commonest among several primary data collection methods are surveys, and they can be grossly categorised into two groups. Viz., manual, and electronic. There are three primary electronic data collection methods: computer-administered surveys, electronic mail surveys, and web surveys. The technology for online survey research is young and evolving. However, many researchers in different disciplines may be unaware of the advantages and disadvantages of conducting survey research online, such as in the public health discipline. Creating a questionnaire, contacting the sample population, storing the responses, visualisation of survey results can be done online. Online surveys are helpful in questionnaire preparation, data collection, storing data, visualisation of data, and collaboration of work. Online surveys can be conducted at a low cost and in a short period. The researcher can start the survey, pause the survey, and restart the survey whenever he wants. Several other studies also stated that online surveys are cost-effective studies and can be conducted quickly. The challenges related to online surveys are the sampling, response rate, non-respondent characteristics, maintenance of confidentiality, and ethical issues. Concluding that, an online survey tool is an internet-based survey tool with advantages and disadvantages in every survey stage. The researcher had to decide to use the online survey tool based on his study setting, study population and methodology of the study.

Keywords: Computer administered surveys, Email based surveys, Online surveys, Survey tools, Web surveys

I. Introduction

Data is of paramount importance for research. Though the methodology for research studies may vary, every research is based on data that should be of good quality and then analysed and interpreted to yield information. Research is predominantly based on primary data. The commonest among several methods of primary data collection is surveys. With the application of probability sampling in the 1930s, surveys became a standard tool for empirical research in social sciences, marketing, and official statistics (1). According to the Oxford dictionary, a survey is an investigation of the opinions or experiences of a group of people based on a series of questions (2). Questions used in the survey are generally predetermined questions, mainly aimed at extracting specific data from a particular group of people regarding their preferences, opinions, behaviour, or factual information, depending on the survey purpose. Surveys can be specific and limited, or they can have more global, widespread goals. Psychologists and sociologists often use surveys to analyse behaviour. At the same time, it is also used to meet the more pragmatic needs of the media, such as in evaluating political candidates, public health officials, professional organisations, and advertising and marketing directors (3).

A single survey comprises at least a sample population, a data collection method, a survey tool with individual questions that become data that would be analysed statistically (3). The success of a research project depends on the population sampled and how representative it is of the target population (3). Data collection methods can be categorised into two groups. Viz., manual, and electronic. In the health sector, data can be collected electronically through electronic health records (EHR). The records contribute to the pool of secondary data. Electronic records provide several advantages over paper-based ones as data is readily available for analysis, such as data mining (4). Surveys have, however, primarily depended on paper-based methods using techniques of household visits, face to face conversations, interviews, distributing questionnaires, filling of schedules under direct supervision etc. The challenge is to transform the paper data into an electronic form for processing and analysis. Even when data is available electronically often, it is reduced to paper format before data capturing. This is inefficient and expensive and results in more inadequate data (5).
The tedious data entry process can be simplified by using Optical Character Recognition (OCR) and Intelligent character Recognition (ICR); however, these can be used only where the survey responses are standardised. With the wide availability of computer systems and internet connectivity, online electronic data collection has several advantages. However, there are several disadvantages too. There are three primary electronic data collection methods computer-administered surveys, electronic mail surveys, and web surveys. (6,7).

The technology for online survey research is young and evolving. Until recently, creating and conducting an online survey was a time-consuming task requiring familiarity with web authoring programs, HTML code, and scripting programs (8). However, survey authoring software packages and online survey services make online survey research much easier and faster. Yet, many researchers in different disciplines may be unaware of the advantages and disadvantages associated with conducting survey research online, such as in the public health discipline. The scope of surveys in the public health field is extensive. Surveys in public health could be for different purposes - evaluating the health status, investigating the factors affecting health and disease, and relating to the administration of health of a population. Simple patient satisfaction surveys to National level Family health surveys, epidemic surveys in a village, to morbidity and mortality surveys are different examples of public health surveys. Sample size, targeted population, questions in the survey form may vary from survey to survey depending on the purpose of the survey. The data collection procedure is mainly offline by household visits, face-to-face conversations, direct observation, etc. However, because of advancements in science and technology, particularly in communication, data collection procedures in the surveys have also changed. The usage of online survey tools gradually started in public health also.

Creating a questionnaire, contacting the sample population, storing the responses, visualisation of survey results could be done online. Some survey tools such as Epi-data, Epi-collect, and Epi-info provide facilities to distribute the survey form to several data collectors for real-time data entry in the software and integration of work done by different data collectors (9–11). They would also provide the facility to store the geo-location data such as longitude and latitude details. They integrate with mobile technology, and the application can be saved as an app in android based or IOS-based cell phones. Despite advancements in the online survey data collection methods, many public health experts are still unaware of the strengths and weaknesses of online data collection. With this background, this paper discusses the advantages and disadvantages of electronic data collection in general and web-based methods in particular.

**Computer Administered Surveys:** With the growth of personal computers and computer networks, computer-administered surveys were the first use of computers to collect survey data. Using a written programme to administer the questions and collect the answers, the survey could be conducted in one of several ways: 1) by gathering respondents to a central location to answer the questions at the computer; 2) the survey installed on the organisation’s network; or 3) the program saved on a disk for respondents to answer and then return the disk (12).

The initial literature compared the data and response quality between computer-administered surveys and paper-based counterparts. Overall, the computer-administered survey showed comparable results to the traditional paper and pencil survey (13–15).

However, electronic surveys were less socially desirable and more extreme (16). Responses to open-ended questions were relatively long and disclosing. No differences were found between the effect of computer-administered surveys, face to face administration, and self-administered paper and pencil surveys on an individual's accurate expression of sensitive information. Though computer administration did not increase accurate response, respondents rated the computer as less friendly but faster, more relaxing, lighter, and more interesting than the other two methods. Better educated individuals and defensive clients did not like using the computer for the survey (17).

Overall the benefits of computer-administered surveys were found to be 1) lowered levels of social desirability responding, 2) shorter and more enjoyable to respondents, 3) elimination of data entry, and 4) use of complex branching and prompting of question (12). Further, computer-administered surveys showed fewer mistakes, blank items and item refusals than paper surveys. There was a decrease in processing costs while still allowing standardisation and anonymity (16,18). Authors have found disadvantages to computer-based surveys. These are 1) expensive for small numbers of people (14) 2) incompatibility of software across systems (19); 3) discomfort in using computers for non-office workers (14) 4) people want a way to know how much time is left.


(19) and 5) resentment to being surveyed in a non-social manner (17,20). Computer administered surveys are most appropriate for organisational settings that allow for 1) a group of people to gather in a central location to answer the questions, or 2) a compatible network of computers to administer the questionnaire.

**Electronic Mail Surveys:** With Email becoming ubiquitous, electronic mail surveys became popular. The survey is sent to a person's email address. The respondents could answer the questions and mail back the responses to the researcher or print the questionnaire fax the responses. Both sending and responding became simplified and cost-effective (12). Other benefits were an increase in the geographical spread of respondents, speed of response- responses can begin immediately, and elimination of time zone hassles for individuals in different geographic areas (21,22). The disadvantages were that data entry was still required and the compatibility of software. The biggest disadvantage was that many people do not have Email, may prefer not to use it, or may be unfamiliar with some of the more advanced functions necessary for answering a questionnaire online. Individuals may delete or not respond to the Email that fails to interest them. Studies have found questionnaire length to be an issue as many programs restrict the number of lines in a message, and respondents will not answer lengthy questionnaires via Email. Krasilovsky, 1996 (23) found that respondents may disregard questionnaire design and rewrite questions, delete questions, and extend scales. Net etiquette (‘netiquette’) frowns upon mass electronic mailings. Email software restricts the number of addresses to which a mail can be sent simultaneously. Over time response rates dropped as consumers (respondents) grew overwhelmed with surveys popping unwelcomed into their inboxes. Companies had to find a way to improve response rates. Survey sites were born, helping match companies to consumers who wanted to share their opinions (24).

**The web survey:** A survey may seem like a simple progression of questions to a survey taker. However, while some surveys can be as short as a single question, others can be complex web blocks of questions and conditions that sometimes would include scripting (25). Modern online survey tools include three main components questionnaire design, distribution, and reporting. Simple Polls are already built into Facebook and Twitter. Google Forms is ideal for sending out a short questionnaire, charting the results or exporting them for analysis to a spreadsheet. It provides for various questions formats from Text boxes, Paragraph texts, Multiple choice, checkboxes, scale, grid etc. It permits custom logic to navigate through questions based on answers. Inbuilt data validation rules ensure that people get the right questions based on the previous answers.

There are several online survey tools for more complex designs for novice and advanced users. Many include display logic, which can show or hide a question or section of a survey based on conditions that occurred before it. Other features include survey reminders, timers, question blocks, kiosk-based responses etc. PC Magazine offers an excellent comparison of the top survey tools (25). However, data collection must be done often in challenging environments with limited or no internet connectivity and power supply. The situation becomes even more challenging during humanitarian crises – natural disasters, civil strife, wars etc. The challenge is to collect data offline using small devices using multiple personnel in the field to be merged later to a single database. KoBo Toolbox, developed by the Harvard Humanitarian Initiative, is an open-source suite of data collection and analysis tools to bridge this gap (26). Data can be collected using Android Phones and other handheld devices. Epi info and Epi-collect offer similar features. A review of 10 software is available at Zapier.com (27).

While email and online surveys have the same advantages and disadvantages as computer-based data collection, many additional challenges are thrown up. The challenges relate to the sampling, response rate, non-respondent characteristics, maintenance of confidentiality, and ethical issues.

**Suitability for different study types:** Online survey technique is suitable for descriptive, case-control, cohort studies and evaluation studies. Online survey technique is suitable to collect information in a cross-sectional study in which the researcher will contact the participant only once and longitudinal studies in which the researcher should contact the participant more times. The web-based data collection procedure is helpful in cross-sectional studies and case-control studies. The researcher can limit the multiple submissions by enabling cookies in the survey tool. The email-based data collection procedure is useful in longitudinal studies and cohort studies. The researcher can send timely reminders to the participant till he gets the response. Online surveys are not suited for experimental studies where direct observations or measurements must be recorded. In Experimental studies, software packages such as Epi-info, Epi-collect are useful. The investigator can enter the details directly in the software package in which the data will be stored electronically. The chief investigator can integrate the work done by different investigators in different locations. Nevertheless, it needs awareness about online technology in the investigators.
Selection of sample population: Probability sampling has been the dominant paradigm for surveys for many decades, but it has by no means been the only paradigm, nor has it always been dominant. The principal goal of survey sampling is to make reliable and accurate inferences to a broader population, often referred to as “representation” (28). Commonly used study designs in health research are case-control studies, clinical trials, evaluation research designs, and intercept surveys are non-probability sampling methods. With the advantages of becoming web-based, surveys have become non-probabilistic because of the state that biases may occur in web surveys due to self-selection, under-coverage, non-response, and sampling errors. Weighting is used to adjust survey statistics for unequal selection probabilities, coverage error or to represent population characteristics on several covariates (29).

In web-based surveys, the researcher cannot ensure the sample population is mainly from the targeted population. Anyone can open the link and fill the form. Kevin B. Wright also stated the advantages and disadvantages of online research are sampling issues, access to a unique population and selection bias of selecting study sample (7). Bobby Duffy and Kate Smith mentioned the same advantages in their study (30,31).

Ethical Issues of Online Survey - Informed Consent form: A basic standard of ethical research is that prospective participants can make an informed choice whether to consent or participate. This is to ensure potential participants are fully informed in online survey research as in other types of research (32). A study conducted by Elizabeth A. Buchanan, Erin E. Hvizdak also mentioned that “Respondents indicated that the electronic and online nature of these survey data challenges traditional research ethics principles such as consent, risk, privacy, anonymity, confidentiality, and autonomy, and adds new methodological complexities surrounding data storage, security, sampling, and survey design” (33). In their study, Connie K. Varnhagen et al. also stated that obtaining informed consent online is not substantially different from obtaining it via paper presentation (34).

As an essential requirement of ethical research, the researcher should seek consent from the participant before taking their opinions. The consent form should be separate, and it should not be combined with the main questionnaire. Mahon (2013) recommends setting an information sheet as the first page of the online survey, with participants required to check a box to indicate consent before accessing the survey. This ensures that participants have access to the same information they receive before completing an offline survey (35).

Page logic option is useful to lead the participant to the main questionnaire. The IP address of the participant will act as a data key to link the consent form and data filled by the participants. However, the Internet assigns a dynamic IP address to each user for each session. If the participant fills the consent form and the data form in different timings, then linking the consent form and data filled by the participant may not be possible. Participants may bypass the information page. Referrer verification can be set up to prevent this, allowing access to the survey only when it comes from a particular URL. Valerie M. Sue, Lois A. Ritter, in their book Conducting Online Surveys, mentioned the same procedure to take informed consent from online survey participants (31).

Privacy, anonymity, and confidentiality are critical ethical considerations in online survey research. Although often considered in tandem, it is essential to differentiate anonymity and confidentiality. Anonymity is the process of not disclosing the identity of a research participant or the author of a particular view or opinion. Confidentiality is the process of not disclosing to other parties opinions or information gathered in the research process (36). The increasing functionality of survey tools can potentially undermine both anonymity and confidentiality.

Andrew Clark (2006) states three broad reasons for anonymising data, the first to ‘protect’ or hide the identity of research participants, the second to disguise the identification of research locations and the third to comply with legal requirements on privacy. The first is particularly important when sensitive, illegal, or confidential information may have been disclosed during the research process or when information is disclosed, which may cause the participant distress should other parties learn such information (36).

IP addresses should be stripped from the dataset, preferably before saving the data file to the researcher’s computer (37,38). Unique tracking links in online surveys also undermine anonymity by providing a link between survey responses and the email address of the survey respondent. These, too, should be stripped from the data set. Clark discusses techniques of anonymising data such as ‘blanket anonymisation’ and the pros
and cons of anonymising data. Direct and, Indirect identifiers need to be anonymised. Several techniques have been described to anonymise data (39–41).

For public health, geo-referenced data is of particular importance for epidemiological analysis, including spatial analysis. Point coordinates may be replaced with non-disclosing features or variables - polygon features (km2 grid, postcode district, county) or linear features (random line, road, river). Alternately the geo-references can be maintained intact and access restrictions imposed on the data (40).

**Questionnaire preparation:** Online survey technique changed the individual questions in the survey tool. The traditional questionnaire will contain only open-ended or structured questionnaires with multiple-choice options. However, various online survey tools provide different options to create a complex questionnaire. These tools also enable us to create automated data, for example, the Date and time of filling the questionnaire, geo-spatial location details of participant, IP address from which the participant is filling the form. Some survey tools, such as Surveymonky.com, will give suggestions in preparing the questions (42). Google forms have many add-ons for integration with other survey tools (43). Researchers can prepare a more complex form by using Yes/No options, Radio buttons, matrix of radio buttons, drop-down menus, matrix of drop-down menus and even can conduct opinion polls. Some survey tools also provide page logic options and question logic options. Based on the answer to the previous question, they will display the next question or forward it to the next page. The progress bar is another feature to display the progress of questionnaire filling. Some survey tools will use cookies to stop multiple submissions from the same IP addresses. A weakness in online surveys in questionnaire preparation is that the researcher cannot probe the participant to get answers and ask leading questions.

**Data collection:** Even though, as stated earlier, an online survey may yield a non-probabilistic sample, a researcher can get diverse participants. Professional groups are available on many websites such as LinkedIn, ResearchGate, WhatsApp groups. A sample population having expertise in one specific field can be obtained from these groups. Kevin B. Wright also stated that the main advantages of online research are access to a unique population, saving time and conducting research at low cost (7). There are some disadvantages also associated with the online survey method. Distribution of survey links is easy, and sometimes it will become viral. The researcher cannot determine questionnaire filling time and participants may abandon the survey giving partial data. Participants can take their own time to fill the form. It may create a bias. It is difficult to explain in detail about study objectives. If the participants have a doubt, the researcher cannot give the answer immediately. Potential Spamming by participants could be another disadvantage. A participant repeatedly submitting the same opinion may alter the results and create bias. The software does have a workaround for this by restricting respondents to one response.

A significant issue of online surveys is the participation rate. Generally, response rates are extremely poor compared to the offline survey method. In our experience, our survey yielded a poor response rate. The survey form was distributed in different platforms such as WhatsApp groups and LinkedIn contacts and mailed to several public health experts. Only 74 submissions were received. Among the 74 total opinions, 14 were duplicate entries, and 60 were original entries. Among the 60 participants, 34 participants gave their opinion, and 26 participants were unwilling to give their opinion. A study conducted by Abhijit Boratne had a response rate of 50% among 76 participants. Only 40% had submitted completed forms. In a survey of 10000 anaesthetists, only 17% responded. A comparison of various survey methods showed an average response rate for all methods to be 33%. In-Person surveys had the highest rates of 50%. Email surveys had 30%, online surveys 29% and in-app surveys 13% (44). Roberts and Allen summarise the reasons for a poor response as faulty email addresses when emails are the primary mode of recruitment, irregular or non-access of emails by potential participants, emails being filtered to the spam folder and survey fatigue, a phenomenon common to all surveys (32). They further state, “In combination, these factors may reduce the response rate to the survey and potentially bias results if unreachable potential participants systematically vary from those who do receive and read the recruitment email”. Researchers recommend that a percentage of non-respondents be contacted, profiled, and reasons for non-response elicited to account for this bias.

The second component of survey non-response is where participants choose not to answer some questions. Internet survey researchers can enable “forced responding”; where a participant cannot move on to a further question until an answer to the current question has been provided. While this has advantages for the researcher in eliminating missing data, it raises ethical concerns. Baker (2012) reported that three-quarters of the 52 IRBs surveyed viewed forced responding as violating research participants’ rights not to answer individual questions [29]
Reliability of the opinion expressed by the participant is another issue. Online survey participants choose midpoints in the scale while filling the forms—the same opinion expressed by Bobby Duffy and Kate Smith in their study [32]. The researcher cannot judge the participants' seriousness about the questionnaire while expressing an opinion. The main disadvantage of the online survey method is missing out on knowledgeable participants lacking computing and internet skills. The small screen size of mobile phones limits the length of the questionnaire and the responses to open-ended questions.

**Storing of Data:** Collected data should be stored in a proper way for statistical analysis. Online survey technique is useful to store data online after submitting the filled-in form. Real-time storing of data is an automatic procedure in the online survey technique. The data will be stored in the central server of the website, and the researcher can download the data from the server. The main threat for online storing of data is the crash of data in servers because of server issues or hacking of the websites. The solution for this threat is to download the data regularly and storing it on a personal computer. Online data storing has an extra advantage, i.e. stored data is not accessible for others; thus, it will be free from data editing.

**Visualisation of data:** Online survey tools will show the data in the form of different charts and graphs. It is very easy to plot the graph between different variables. Bobby Duffy and Kate Smith also stated that online survey tools are flexible to visualise the data (7,30). However, the main disadvantage is that the survey tools will have only some prefixed models. Analysis of data is not possible in online tools. However, this deficiency can be overcome by using software packages such as Epi-info, Epi-data or Epi-Collect.

**Collaboration of research:** Online survey technique is useful for multicentric studies. Two collaborators from different areas can share the work, share the data and results. But the disadvantage in sharing the project is, two researchers can modify the questionnaire without the knowledge of another researcher. It can be controlled by regular communication between collaborators. Sharing data and results may raise privacy issues for study participants. There may be a threat of data leak, and data security is another problem when multiple researchers are involved.

**Other advantages:** Online surveys can be conducted at a low cost and in a short period of time. The researcher can start the survey, able to pause the survey and restart the survey whenever he wants. Several other studies also stated that online surveys are cost-effective studies and can be conducted in a short period.

### II. Conclusion

An online survey tool is an internet-based survey tool with advantages and disadvantages in every survey stage. The researcher has to decide to use the online survey tool based on his study setting, study population and methodology of the study.

### References


Strengths and Weakness of Online Surveys

[10]. Epicollect5 - Mobile & Web Application for free and easy data collection. [Internet]. [cited 2018 Nov 27]. Available from: https://five.epicollect.net/


DOI: 10.9790/0837-2405053138 www.iosrjournals.org 37 | Page
Strengths and Weakness of Online Surveys


