Hooley, Tristram, John Marriott, and Jane Wellens. "Online surveys." *What is Online Research?: Using the Internet for Social Science Research*. London: Bloomsbury Academic, 2012. 39–52. The 'What is?' Research Methods Series. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781849665544.ch-004>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 11:18 UTC.

Copyright © Tristram Hooley, John Marriott and Jane Wellens 2012. You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

4 Online surveys

The use of online questionnaires to gather survey data has become ubiquitous. They offer huge value to researchers in terms of cost, speed of data collection and analysis, and access to respondents. Given these advantages it is unsurprising that their use extends far beyond the world of academic research. Web users are likely to encounter online surveys frequently, and many of these will have been issued by service providers or marketing companies rather than academic researchers. This experience may lead some web users to consider most online surveys as little better than spam. For this reason academic researchers have to ensure that their surveys exemplify good practice; they should be aware that engaging a target population in an online survey may not be a simple or straightforward task. This chapter explores some of the issues surrounding online surveys, provides readers with an overview of the method, and a range of practical advice on how to develop effective surveys, maximize response rates and minimize dropout rates and missing data.

case study

Using online surveys to investigate childhood cancer survivors

Cantrell and Lupinacci undertook research with early childhood cancer survivors in order to examine how the experience of cancer had impacted on their physical and psychosocial characteristics, and to explore their quality of life (Cantrell and Lupinacci 2007; Cantrell and Lupinacci 2008). They chose to undertake the study using an online survey and they discuss a range of reasons for doing this – such as cost-effectiveness, decreased data collection time (compared to other methods) and increased accuracy of data collection. However, they argue that the methodology was particularly appropriate because it enabled them to reach a ‘hidden’ population who may not wish

to identify themselves openly as cancer survivors. Furthermore, the respondents were geographically dispersed and hard-to-reach using conventional survey techniques, as there was no central database of the population. Because their target population was young adults (aged 22–28), the researchers were able to assume a familiarity with using the web.

Cantrell and Lupinacci developed a pilot online questionnaire and trialled it with ten healthy adults. Individuals who participated in the pilot were asked to say how long the questionnaire took to complete and to make suggestions to improve usability. Following the pilot, some modifications were made to the survey. It was then launched and advertised on six websites used by childhood survivors of cancer. The articles posted on these websites included an introduction to the research protocol, a pledge of anonymity, the researchers' email addresses, information about the investigators' research and clinical nursing experience background, and a direct link to the online questionnaire.

The questionnaire was designed to take around 15 minutes to complete and included 167 questions, which is arguably more than would be ideal. Participants were required to complete every question to try to reduce missing data. However, after 16 responses the researchers recognized that a long questionnaire entirely comprising mandatory questions was likely to lead to increased dropout rates. The researchers chose to relax this requirement for all subsequent participants, allowing them to skip questions as they chose. Cantrell and Lupinacci considered whether the missing data could have been minimized by making fewer questions optional or by changing the nature of some of the questions that were asked. However, the study was dealing with a sensitive topic and they reflected that if some questions were mandatory they might have resulted in emotional distress to some participants.

This case-study (as with others in this book) is not advanced as a flawless exemplar of good practice, but rather shown as a concrete example from which researchers can learn and adapt for their own purposes. Indeed Cantrell and Lupinacci are largely concerned with considering what methodological learning can be extracted from their experience (Cantrell and Lupinacci 2007). Furthermore, in 2008 Larew constructively critiqued a number of areas of Cantrell and Lupianacci's

methodology, including their approach to sampling and the decision to make questions compulsory, as well as reflecting on how changes in technology open up further possibilities for enhanced survey design (Larew 2008; Cantrell and Lupinacci 2007). However, Cantrell and Lupinacci's experience remains interesting, particularly because it illustrates the way in which the study's methodology was built out of the research questions and closely related to the needs of the target population. Furthermore, the willingness to pilot, adapt and reflect on the methodology during and after the study demonstrates a critical reflexivity that researchers working with the constantly shifting world of online research methods are likely to need.

Survey design

Whilst this chapter focuses on using survey methodologies online, it is important to recognize that there are strong continuities between online and onsite research. Most of the basic principles of survey-based methodologies remain the same when they move online. Rather than discuss the basics of survey design, this chapter focuses on how these issues are recontextualized by moving the survey methodology online. Those new to survey-based methods might want to read this chapter alongside a more general book such as *Survey Research Methods* (Babbie 1990); *Conducting Survey Research in the Social Sciences* (Newman and McNeil 1998); *Survey Research Methods (Applied Social Research Methods)* (Fowler 2008). Any of these books (alongside many others) will provide the reader with a solid introduction to methods of survey-based data collection and analysis.

It is perhaps worth beginning by providing definitions for some key terms that will be used in this chapter:

- **Surveys:** research methods which researchers can use to collect data. They can be either quantitative or qualitative in focus, or a mixture of the two.
- **Questionnaires:** tools or instruments which researchers use to undertake a survey. They usually comprise a series of questions or stimuli for response.
- **Populations:** the total group of people that are being studied. Researchers will commonly be unable to interact with everyone

being studied (the population), and will therefore need to work with a sample.

- **Samples:** part of a population that is examined for the purpose of drawing inferences about the population as a whole. Quantitative researchers may use statistical techniques to determine their sample and to analyse the data they gather.
- **Recruitment:** the ways in which participants are encouraged to take part in a survey.
- **Response rates:** the number of people who participate in a survey in relation to the number of people in the sample. This is often expressed as a percentage.

These definitions are basic, but it is recognized that online research methods cut across a range of disciplinary and methodological backgrounds, so it is useful to establish some core definitions. These definitions can be found alongside a range of other terms in the glossary associated with this book.

When to use an online survey?

Developing the appropriate methodology(ies) to address a particular set of research questions is one of the most critical decisions that a researcher makes. If a survey-based methodology is determined to be suitable, the researcher then has a range of decisions to make about the most appropriate way to collect data, engage participants and manage the survey. Surveys can be conducted face-to-face, by post, by telephone and by a range of other mechanisms. The decision to use an online survey must therefore be considered carefully and weighed up against the alternatives. The purpose of this chapter is not to encourage researchers to use online surveys regardless of context, but rather to help guide their appropriate use.

Despite sounding this note of caution, online surveys are becoming more frequent compared to alternative survey methodologies and this is likely to continue as the demographic penetration of the internet increases. It is therefore useful to consider some of the advantages and disadvantages of online surveys. Madge *et al.* argue that online surveys have the following advantages (Madge, O'Connor, Wellens, Hooley and Shaw 2006):

- Speed and volume of data collection (Fleming and Bowden 2009).
- Savings in costs.
- Flexible design.
- Data accuracy.
- Access to research populations.
- Anonymity.
- Respondent acceptability.

Many of these advantages are discussed in the Cantrell and Lupinacci case-study (Cantrell and Lupinacci 2007; Cantrell and Lupinacci 2008). However, online surveys are clearly not appropriate in all circumstances and with all populations. Furthermore, there are a number of downsides associated with online surveys. These are also discussed by Madge *et al.*, who note the following:

- Sample bias.
- Measurement error.
- Non-response bias.
- Length, response and dropout rates.
- Technical problems.
- Ethical issues.

This chapter returns to some of these issues in more detail later on. The key sampling issues are, however, whether the population that can be accessed using an online survey is different to that which can be reached using other survey approaches (sample bias), and whether respondents behave in a different way because they are participating online as compared to another survey method (measurement error). There is a range of potential technical and ethical issues that are associated with the online environment (discussed in more depth in chapters 2 and 3). However, many of the advantages and disadvantages associated with online surveys are both contested and dynamic. As discussed in chapters 1 and 2, the online environment continues to change and develop as more users are attracted to it, technologies change and as the cultural position of the internet itself shifts. Researchers are likely to need to continue to innovate in the methodologies that they use in order to continue to address and respond to these changes. There are no rules about when an online survey should be used as opposed to some other form of survey. The task is, as ever, to use methodologies that are appropriate for the

research questions and populations that you are examining. Obviously if most of your population are not internet users then an online method is clearly inappropriate, but most of the time the issues are likely to present in more subtle ways. For example, if the verification of identity is an important aspect of your research then a conventional online survey might present problems.

Discussing the phenomenon of online identity manipulation, Roberts and Parks question whether we can trust the identities that people convey through online personas (Roberts and Parks 2001). However, the idea that an offline identity exists that is more real than an individual's online identity has been frequently challenged (Turkle 1999; Bowker and Tuffin 2003; Valkenburg and Peter 2008). A related issue is whether researchers are more able to judge the trustworthiness of participants in a face-to-face environment. Hewson *et al.* argue that onsite a researcher is able to pick up on a variety of visual and social clues to judge the trustworthiness of the response (Hewson, Yule, Laurent and Vogel 2003). However, research around online identity suggests that different but equivalent clues exist to signify authenticity, identity and trustworthiness (Toma and Hancock 2009; Mislove, Viswanath, Gummadi and Druschel 2010; Shin and Kim 2010). So, even if a research question requires a mechanism for identity verification and assessing trustworthiness, it does not necessarily answer the question as to whether to use online or onsite methods. In the case of an online survey, identity might be verified through the development of a password-protected site and the issuing of passwords. However, the use of passwords does not 'solve' the problem of identity verification any more than physically meeting someone guarantees their truthfulness, but careful methodological thinking can aid the creation of an appropriate tool for data gathering and put in place processes that minimize the likelihood of receiving problematic responses to surveys.

The question of whether to use online or onsite methodologies is also becoming increasingly blurred with the growth in power and availability of mobile devices which make it possible to blend onsite and online approaches together. Face-to-face surveys are often conducted using a tablet PC, with results being sent to a database during the course of the interaction. This makes it increasingly possible to blend together the results that are generated through onsite and online surveys, but what are the methodological issues in doing this? Converse *et al.* undertook a mixed postal-/web-based survey and reported increased response

rates (Converse, Wolfe, Huang and Oswald 2008); this was confirmed by Dillman *et al.* who combined together a wider range of different survey methods (Dillman, Phelps, Tortora, Swift, Kohrell, Berck and Messer 2009). However, Sax *et al.* and Dillman *et al.* also concluded that there were some differences in the way people responded to the different modes of survey (Sax, Gilmartin and Bryant 2003). How people respond to online surveys in comparison to other kinds of survey has not been comprehensively mapped and is likely to remain dynamic. However, some research argues that web-based surveys have lower response rates (Shih and Fan 2008), but deliver less missing data (Dolnicar, Laesser and Matus 2009). In contrast to this, several studies which have investigated this issue have found that there are few differences in the way people respond to different survey modes (Arnau, Thompson and Cook 2001; Gosling, Vazire, Srivastava and John 2004; Fleming and Bowden 2009).

Online surveys and instrument design

Following the decision to utilize an online survey, the next step is to design an instrument that can be used to collect data. There are a number of challenges which need to be considered whilst developing an online survey, and Couper provides a comprehensive overview (Couper 2008).

Issues associated with online survey design can broadly be described as relating to either the questionnaire's *accessibility* or its *usability*. Accessibility describes elements of the design that either prevent, limit or enable the whole population responding to a survey. Accessibility issues commonly relate to disability or to the challenges of responding to a survey from a range of different platforms with different technical specifications. Usability relates to the design features that the respondent encounters whilst they are trying to assess and respond to the survey.

Poorly designed websites can pose considerable barriers for individuals with disabilities. Well-designed sites take account of the fact that people might be accessing the site using screen readers, or that they may seek to change background colour or increase font size to enhance readability. There is an extensive literature on web accessibility, and it is probably worth consulting a basic introductory text such as the W3C's accessibility materials (W3C 2010). Accessibility is also an important consideration when thinking about the selection of an appropriate tool to use for your research. The issues of ensuring access to online surveys for people with

disabilities often require similar solutions to more general issues about accessibility relating to different operating systems, browsers and devices.

There are tensions between the functionality and the accessibility of all resources on the web. Broadly speaking, the more things a website/online survey can do, the more problems that people will encounter in accessing it. People access websites from a wide range of different technical environments, and the variations that exist in browsers, operating systems, screen size and bandwidth will mean that the experience of the designer of the survey is unlikely to be the same as the experience of the respondent. During the survey design this variability can be explored through piloting in a variety of environments and enabling flexibility and user configurability. Crawford, McCabe and Pope recognized this and called for the creation of survey design standards (Crawford, McCabe and Pope 2005). They argued that there were four main areas in which researchers might want to establish clearer standardization: screen design; questionnaire writing; respondent communication and process standards. This attempt moves beyond merely ensuring accessibility, and towards developing enhanced usability for respondents to web questionnaires.

Despite the arguments made by Crawford, McCabe and Pope, it is difficult to fully standardize the area of web usability. Individuals' preferences in the use of web resources vary, and thus specifying design features which facilitate increased engagement can be challenging to render as a set of standards. There is a considerable body of research and practice in the area of enhancing web usability. Offering a practical introduction to this area, Krug emphasises the importance of paying attention to existing conventions (what do everyone else's online surveys look like?) and undertaking formative piloting of the instrument during its development (Krug 2005). But there are a number of established areas to pay attention to in the design of online surveys:

- **Transparency:** Respondents should be able to easily ascertain the purpose, rationale and context of any survey they participate in. In practice this means that background information such as who is undertaking the survey, who is funding it, what the main research questions are, how data will be stored and where findings will be reported should all be available to respondents. A short welcome screen providing a summary of the project, with an associated box to indicate consent, is also a key element of this.

- **Consistency:** Online questionnaires should be designed to be internally consistent. Respondents are likely to find changes in font, colour and navigation disorientating. Key to this is the importance of limiting the range of question types used across a questionnaire. Each question type that is introduced (such as multiple choice, multiple answer, ranking) presents respondents with a new usability challenge. Further discussion of different question types can be found in Dillman, Tortora and Bowker, Best and Krueger, and Healey. (Dillman, Tortora and Bowker 1998; Best and Krueger 2004; Healey 2009). In general, choose simple question types which users are able to parse in a single glance and avoid the proliferation of question types. Crucially, there is value in piloting a questionnaire before launch and watching someone completing it. Observation of this kind is a very powerful tool for identifying which questions are likely to pose usability challenges.
- **Brevity:** The flexibility and low cost of the online form allows researchers the potential to create very long questionnaires. However, many researchers argue that shorter online questionnaires receive higher response rates and less dropout than longer ones (Deutskens, de Ruyter, Wetzels and Oosterveld 2004; Ganassali 2008; Galesic and Bosnjak 2009). Given this, there are clearly strong reasons to try and ensure that questionnaires are kept short and that appropriate use is made of routing questions to ensure that respondents are answering only the questions that they need to. Dillman, Tortora and Bowker argue that the inclusion of a progress indicator to inform respondents about how much of the survey they have completed will reduce dropout rates. However these positive effects are likely to be realized only when the progress bar is an accurate indicator of the amount of time taken (Yan, Conrad, Tourangeau and Couper 2010).
- **Respondent autonomy:** Respondents to online questionnaires are usually giving their time for free and without any promise of personal advantage. Given this, it is important that the experience is a positive one. Furthermore, researchers have a duty to conduct research in an ethical way without compulsion. In practice this means that making questions compulsory should be avoided where possible. In addition, it should usually be possible to exit the questionnaire whilst ensuring that the data you have submitted is included in the survey. These issues of respondent autonomy clearly need to be balanced

with concerns about missing data. However, it is also important to recognize that respondent desire for autonomy is likely to increase where sensitive subjects are being dealt with.

Choosing a tool to deliver the survey

There is a vast array of different tools that deliver online surveys. In the early days of using online surveys, researchers often had to struggle to access online survey tools or to build their own. Madge *et al.* include a detailed technical guide that takes researchers through the process of crafting an online survey from scratch (Madge, O'Connor, Wellens, Hooley and Shaw 2006). Increasingly, however, it is unnecessary to build a bespoke survey tool since there is a growing range of purpose-built tools which researchers with basic IT literacy can use effectively. It is not possible to provide a list of such tools because the market is shifting quickly – new tools are developed regularly and old ones fall out of use. A judicious use of a search engine using 'online survey tools' will generate numerous options for you to investigate, many of which will be free or low cost, and it can be useful to look for review sites to help in your decision making. There is also a number of sources that will provide more academic discussions of survey tool selection (Wright 2005; Madge, O'Connor, Wellens, Hooley and Shaw 2006; Kacmirek 2008). If you have access to technical support (for example, through a university) it may be useful to explore and consider any options that are actively supported. However, in choosing a survey tool it may be useful to consider some of the following issues:

- **Accessibility and usability:** These issues have already been discussed above, but it is important that they are attended to at the point at which a tool is being selected.
- **Analysis tools:** Does the tool offer any analysis functionality? This may not be necessary if you are likely to move the data into another format for analysis, but such functionality can be useful to gain a quick overview of progress whilst the survey is ongoing.
- **Configurability:** How much control is it possible to exert over the way that a particular tool looks/operates. This may relate to basic issues such as colour and font, but is also likely to be critical when looking at issues such as page design and routing of questions.

- **Data security:** You have ethical, and possibly legal, responsibilities to ensure that your data is kept in a secure fashion. Examining the data security of any tool that you select is therefore crucial.
- **Ease of use/support:** There are big advantages to choosing a tool that you can use easily yourself without the requirement for technical support.
- **Export formats:** Does the tool allow you to export data into the format in which you will analyse it – for example, SPSS or Excel?
- **Question type:** What range of questions do different tools offer, and do these question type match with your needs?
- **Vendor limits:** Some vendors place limits on things such as the maximum number of questions, amount of responses or file size available. Often these limits can be removed by paying for an advanced license. However, it is important to investigate the limits that are being placed and match them to your needs.
- **Verification tools:** Does the tool offer any way to verify identity or manage spam? There are a variety of approaches, including the issuing of passwords or the verification against email addresses or other online accounts such as Google or Facebook.

Once a tool has been chosen it is recommended you pilot it to ensure that it is able to deliver all of the functionality that is required.

Sampling and recruitment

For many researchers, understanding sampling issues and developing appropriate recruitment strategies is likely to be key to the successful adoption of online methods. Some issues relating to sampling and recruitment have already been discussed, including the need to craft carefully the research questions and select the target population, as have the issues around verifying identity. However, just as problematic as verifying the identity of a particular respondent is the difficulty of establishing a clear picture of the population that can, or might, view the survey. For researchers coming from a quantitative tradition, this lack of clarity can present serious difficulties. However, as with many issues relating to online research methods these issues are not unique to the online environment, they are just exacerbated by it. It has always been possible to conduct research on 'broad and diffuse populations' (Couper 2007: 588), and it

has always been challenging to construct sampling strategies in relation to these groups. One of the effects of moving these kinds of surveys online is that any attempt to recruit widely runs the risk of exceeding the population altogether by, for example, engaging respondents from across the world in research that was designed for a particular national context. However, Norman and Russell argue that this 'pass-along effect' should not always be seen as negative and may, with careful consideration, intentionally be incorporated into the sampling strategy (Norman and Russell 2006). It is important to recognize that the ability to access a wide population through the dissemination of surveys via online networks also offers considerable advantages, such as making traditionally hard-to-reach groups more accessible (Dodd 1998: 63). Furthermore, it is also possible to use statistical strategies such as weighting to reduce bias in some cases (Best and Krueger 2004).

The problem with defining a web population is that there is no central database of online identities of the kind that is provided (in the UK and many other countries) by the electoral roll. In addition, it is frequently difficult to link what online information does exist to other information about an individual, such as their geographical position or demographics, although (as is discussed in chapter 8) this is becoming increasingly possible. In certain contexts, such as surveys of single organizations or networks, it is possible that, with the collaboration of a gatekeeper, researchers will be able to gain a clearer picture of the population. In these cases a wide range of more conventional sampling approaches and statistical techniques becomes useful. In fact, where survey responses can be joined up with existing data that is held on individuals, either by the researcher or on the internet, a range of analytical possibilities are opened up. However, the process of mashing datasets together is ethically challenging, especially if some of this data is harvested from the internet without the permission of the individual. In particular, it is likely to challenge anonymity unless very carefully handled (O'Hara and Shadbolt 2008).

Regardless of the approach that is taken in constructing a sample frame, researchers are likely to experience challenges in recruitment and ensuring a good response rate. One useful element of conducting an online survey is that normally the turnaround time between sending out the survey and receiving responses is relatively short. Crawford, Couper and Lamias suggest that if people are going to complete a web-survey they will do so in the first few hours or days of receiving it (Crawford, Couper and

Lamias 2001). However, it has also been found that increasing response rates can be achieved by follow-up reminders (Schaefer and Dillman 1998; Crawford, Couper and Lamias, 2001). Response rates can also be improved with introductory letters or emails (Porter and Whitcomb 2003), by recommendation from a trusted intermediary (Fang, Shao and Lan 2009), or by increasing the social presence of the researchers in the survey (Best and Krueger 2004). Madge *et al.* have constructed a checklist based on a synthesis of the literature around response rates to online surveys (Madge, O'Connor, Wellens, Hooley and Shaw 2006).

Improve response rates: a checklist (Madge et al.)

- 1 Send introductory letter outlining project and estimated time needed to complete the questionnaire.
- 2 Include an institutional/official website to help validate researchers' identity.
- 3 Provide clear instructions on how to complete the questionnaire.
- 4 Request personal information at the start of the questionnaire rather than the end.
- 5 Use simple questionnaire format and avoid unnecessary graphics.
- 6 Avoid grid questions, open-ended questions and requests for email addresses.
- 7 Design the survey so that it takes approximately 10 minutes to complete.
- 8 Do not include more than 15 questions.
- 9 Send one or two follow-up reminders.
- 10 Include 'social presence' (information to increase trust in the researchers) or missing data messages (thanking participants for completing the survey and informing them about their progress within it) to reduce item non-response.
- 11 Emphasise confidentiality.

In summary

Online surveys offer researchers a powerful tool with many advantages. It is anticipated that as the social penetration of the internet increases, and online and mobile technologies continue to become more embedded

in the lives of the majority of people in the developed world and to a sizable global minority, the utility of online surveys will continue to grow. However, the use of online surveys is not without challenges in areas such as identity verification, sampling and representativeness. For those working in the quantitative tradition, these issues may pose considerable concerns. Thankfully there is a lively and ever growing research literature which is examining these issues and developing strategies to address them. In general, as with other online research methods, online surveys should not be viewed as a panacea. Rather, online surveys need to be used carefully and critically and combined with onsite and postal approaches where appropriate.

Further reading

There has been a considerable amount of energy devoted to the discussion and exposition of online survey-based methods. *Designing Effective Web Surveys* (Couper 2008) provides a good overview of the area, as does *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (Dillman, Smyth and Christian 2008). A very useful practical guide is *Successful Online Surveys* (Healey 2009), which is available online for free, while *Don't Make Me Think* (Krug 2005) provides a solid grounding in the web-usability issues that are central to effective online survey design.