Keywords: *Data, Structured Interview, Self-completion, Web collection.*

Introduction

This fact sheet will discuss the main methods of survey data collection. It will focus on the actual practical task of getting data from your respondents, assuming that you have already learned the basics of finding your sample and writing your collection instrument.

Of course, it is not just those that are collecting data that need to understand the process. Secondary analysts who are looking at data produced by someone else must ensure that they understand full the way in which their data were collected so that they are aware of any potential bias and error.

The data collection phase of the survey process is absolutely vital and not only needs careful planning, but careful monitoring as well. If the data is to be of a high quality the researcher must ensure that they do everything in their power to maximise the response rates, ensure that interviewers are collecting the data correctly and checking the data coming in for inconsistency and error.

Data Collection

How you choose your data collection method depends on a number of quite complex factors, as Floyd and Fowler state:

"The choice of data collection mode – mail, internet, personal interview or group administration – is related to the sample frame, research topic, characteristics of the sample and available staff and facilities; it has implications for response rates, question form and survey costs" (Floyd and Fowler 2002: 58)

Of course your data collection strategy depends on other factors as well, such as the amount of time that you have to collect it, money available in your budget and the complexity or nature of the questions.

Punch lists four key points to think about before you start to design the data collection phase of your survey:
1) *Better planning for data collection, including a professional and ethical access and approach to the respondents, leads to a better quality of data.*

2) *Greater effort in data collection means better quality of data.*

3) *More researcher control in data collection means better quality data.*

4) *If the choice arises when considering data collection strategies, it is better to have a smaller body of good data than a larger body of poor data.*

(Taken from Punch 2003: 41)

The rest of this fact sheet will detail four main methods of survey data collection:

1) Face to Face Interviews

2) Telephone Interviews

3) Mail / Self Completion

4) Internet Collection

Key readings and links follow at the end. This is not a comprehensive introduction to the subject, there are many more books available and it is strongly recommended that you use this as a starting point for your study.

The following table sets out the main components of each method of collection and their advantages and disadvantages.

	Face to Face	Telephone	Mail / Self Comp	Web
Sample	Post Code Address File	Random Digit Dialling	Postcode Address File or local group	
Sample Type	Probability	Probability	Probability	Non Probability
Turn around time	Slow	Fast	Fast	Fast
Cost	High	Medium	Low	Very Low
Interviewers Required	Yes	Yes	No	No
Interview Length	Up to 2 hours	Maximum ½ hour	Maximum 15 minutes	Maximum 15 minutes
Response rates	High	Medium	Low	Low
Main Advantages:	High response rates Better quality of data More complex questions Longer time to interview hence more data collected Interviewer rapport with the respondent	Low cost Able to reach a large number of geographically spread population Fast turnaround time	Low cost Able to reach a large number of geographically spread population	Low cost Able to reach a large number of geographically spread population Able to use visual aids in web surveys
Main Disadvantages	High set up costs Interviewers need training and supervision Long time in the field	Low response rates Sampling problems with key groups Unable to ask long or complex question or use visual aids	Low response rate Poor quality of data if respondents misunderstand the questions No control over respondent selection	Sampling issues Poor quality of data Low response rates

Face to Face Interviews

Almost all of the large scale UK social surveys represented on the Question bank use this method of data collection, which is seen as the gold standard of survey interviewing.

In face to face data collection a team of interviewers specifically trained to collect data for that survey are sent out to respondents homes to administer the questionnaire. This is preceded by an 'advance letter' sent out by the survey organisation to let the respondent know that they have been selected for the survey, what the survey is about, why they have been selected and that an interviewer will be calling.

The covering letter is used in postal surveys as well as face to face data collection, I will treat them as one letter here designed to introduce the survey to the respondent whether there is a postal questionnaire or information that an interviewer will call. The advance letter is crucial to the response rates for the survey. It must be well designed, explain the survey purpose, the organisation behind the survey and why the respondents view is needed. It must also reassure the respondent that all data will be confidential and used only for the purposes of that survey. Some survey agencies have used incentives such as stamps which are included with the letter. This is to help the respondent remember the letter so that when an interviewer calls and asks them if they received a letter they can also say 'you may remember that it included a book of stamps'.

After the respondent has received the advance letter the fieldwork period can begin. A team of interviewers will be 'issued' with sampled addresses in their area. Each interviewer will call at their addresses and follow instructions on how to select a respondent in that household. If there is someone at home and they are both suitable and willing to take part then the interviewer goes ahead and administers the questionnaire. If the interviewer gets no response they may call back at that address up to 8 times at different times of the day to try to get a respondent.

Approaching the respondent and getting agreement for an interviewer is not an easy task. The interviewer must be confident, have an appearance that suits the neighborhood that the respondent is from and be clear about why they are at the doorstep. In practice the survey agencies in the UK such as the Office for National Statistics and the National Centre for Social Research have a field force of interviewers which they use for all of their surveys. These interviewers are trained in general interview techniques and then before fieldwork begins they are briefed on that particular survey. Interviewers are based across the UK and only come together for the survey briefing. Interviewer briefing days normally consist of the survey manager detailing the survey and why it is being carried out. If the survey is carried out annually and the interviewers worked on it before then key results will be fed back. Then there is a comprehensive run through of the questionnaire which is usually loaded on to a laptop given to the interviewer. The briefing days allow the interviewers a chance to talk to other interviewers and discuss issues related to their work. Once the fieldwork begins the interviewer works alone with the support of a Supervisor to collect their data. The interviewers task is not as easy as simply getting a respondent and collecting data. Much of an interviewers time will be spend travelling to addresses, chasing respondents who are not at home, dealing with entry phones and other obstacles and then downloading the data to the field centre. Sapsford estimates that *"a single hard working interviewer is unlikely to manage more than 40 half hour interviews in a week, and fewer would be reasonable"* (Sapsford 1999:109)

Almost all of the face to face interviews carried out in the UK today, especially those from the large survey agencies , are carried out using a laptop in what is known as Computer Assisted Personal Interviewing. The Qb has a separate fact sheet on CAPI.

Face to face interviews are seen as the best method of obtaining data from respondents in large scale surveys. Data collected can be complex and rich, however this method is very expensive.

Telephone Data Collection

Structured interviews using the telephone became popular in the United States in the 1960s as a result of a very large spread of the population and advances in telephone survey methods. Telephone surveys are less popular in the UK and are mainly used for market research, or for following up a face to face interview. Sometimes when a respondent cannot be reached for a face to face interview as a last resort they will be contacted by telephone and a shorter, basic questionnaire will be administered, this is called a Proxy telephone interview.

In a telephone survey the interviewer sits in a call centre and reads a structured questionnaire to the respondent keying in their answers to a computer screen. Computer Assisted Telephone Interviewing as this is known, allows the researchers to have more complex questions, although the questionnaire must be kept shorter on the telephone than in face to face encounters. In a telephone survey questions need to be short and are generally attitudinal or behavioral this is because people are much less inclined to answer personal questions over the telephone when they have not met the interviewer. Response categories also need to be short as respondents will not be able to see show cards and are unlikely to remember a long list of responses read out to them. Generally a telephone questionnaire cannot last for more than half an hour. (Czaja and Blair 1996)

A number of issues need to be considered for a telephone interview to be successful. In the UK today a number of people, especially young people living away from home are likely to have a mobile phone rather than a landline. Many people are not listed in the telephone directory making sampling difficult. Once a respondent is selected and contacted the interviewer may not know what situation they are in at the time, for example if there is another person in the room which may affect the responses.

Mail Surveys / Self Completion Data Collection

Mail surveys or postal questionnaires which are filled in by respondents and then sent back to the researcher are a relatively cheap method of surveying a large sample, especially if that sample is widely geographically dispersed. Generally a questionnaire is sent out with a covering letter and a stamped addressed envelope. Follow up postcards can be sent several times to boost the response rate. Self completion questionnaires can also be given out to certain groups of people at key events such as students leaving a lecture or fans leaving a football match.

Some surveys also use self completion diaries as a way of collecting information from respondents over a certain time period, usually one or two weeks. The UK Expenditure and Food Survey for example, uses diaries over a two week period to collect information on each item of food and drink purchased by the respondent and their families. The use of diaries is clearly a larger burden for respondents who have to fill them in each day although they do have the clear advantage that a great deal of data can be collected which may not have been recollected by the respondent during a survey interview.

There are a number of important considerations when using the self completion method. The only contact the respondent will have with the researcher is through the paper instrument and so the letter accompanying the questionnaire and the questionnaire itself must be well designed.

Firstly, the design of the instrument is crucial as respondents will be filling it in without any guidance from the researcher, questions and instructions must be clear. Details are of crucial importance such as the size of the font and the colour of the paper (Alreck and Settle 1995). As there will be no ability for prompting of answers the coding categories must be well thought out and include all possible answers. This means that there is little scope for including complex questions or a lot of routing instructions as respondents may get confused and not complete the questionnaire. The questionnaire should not take more then ten to fifteen minutes to complete.

Secondly, the researcher should have spent time at the design stage testing the questionnaire with respondents similar in characteristics to the target sample. There may be a number of issues such as a language barriers, use of slang or a low literacy level that needs to be factored into the design. Where possible, questions should be targeted at the appropriate level for the sampled population. One problem with this method of collection is that it is impossible for the researcher to know who in the household or group filled in the information and if they were the most appropriate person to do so.

Thirdly, response rates in self completion questionnaires are often low and this is a major disadvantage of the method. When the questionnaire is send or handed out and the researcher is asking for it to be posted back a self addressed envelope should be included and the address added onto the actual questionnaire. Response rates depend on the interest of the respondents in the subject of the survey and whether or not they think it affects them.

Web Surveys and E-mail Data Collection

Online data collection is growing in popularity as more and more people have access to the internet. It is cheap, fast and does not reply on the traditional geographical boundaries allowing international collection. There are two ways of collecting data using the internet: e-mail and online surveys.

E-mail data collection is fairly self explanatory. The respondent receives the questionnaire either embedded in the e-mail or as an attachment and then sends it back by hitting reply. An embedded questionnaire is often easier to send back as there is no opening of documents, saving and reattaching, plus the fact that respondents may be wary of opening attachments and getting a virus. Respondents may have very different computer software to the researcher and so the e-mail needs to be as plain as possible as not all e-mail software will be able to interpret complex texts, images or other embedded information. This can make an e-mail questionnaire look plain and unwieldy to the respondent. For this reason e-mail questionnaires need to be short and simple.

Web surveys are constructed using software which creates a form for respondents to fill in. There are many different types of software to accomplish this and the researcher needs some degree of knowledge to be able to design and create the online questionnaire. Once the questionnaire is ready careful thought needs to be given as to how to get it to the respondents, will it be sent as a link in an e-mail or features as a pop-up window on a web site? Both of these have problems, as people often delete e-mail they think of as not directly relevant to them and modern browsers often block pop-up windows automatically. As well as this thought needs to be given as to where the responses are going to be sent and kept. Surveys that have a submit button at the end usually have code embedded in them that sends the responses to a server where the data is stored. This server needs to be secure for confidentiality reasons and the researcher will need to access it frequently to check the data.

Web surveys do have a number of important advantages over e-mail collection. The survey form can be much embellished making it more interesting for the respondent. Skips and routing can be built into the form so that respondents do not have to follow tedious instructions. The data can be downloaded straight into a data analysis package saving time coding and typing in responses.

There are two major problems with this type of data collection: sampling and response rates. If you are conducting a survey of say all students in your college or university then the sampling frame of all e-mail addresses will be easily available. However if you are conducting a sample of the more general population then sampling e-mail addresses is very difficult to achieve. Although internet use is growing, not all members of the population have access to e-mail and the internet and are therefore automatically excluded. The response rates to such surveys are notoriously low with the added problem of the researcher not knowing who exactly has responded, where they are and what their motives were.

Further reading and References

Key texts:
Bryman, A (2004) *Social Research Methods* Oxford University Press, Oxford
Sapsford, R (1999) *Survey Research* Sage, London

Bibliograhpy / Referneces:
Corbetta, P (2003) *Social Research: Theory, Methods and Techniques* Sage, London
Czaja, R and Blair, J (1996) *Designing Surveys: A guide to Decisions and Procedures*, Pine Forge Press, Thousand Oaks, California
Dillman, D (2000) *Mail and Internet Surveys: The Tailored Design Method* John Wiley and Sons, London
Floyd, J and Fowler, JR (2002) *Survey Research Methods* 3rd Edition, Sage, London
Groves, R et al (2004) *Survey Methodology* Wiley, London
Punch, K (2003) *Survey Research: The Basics* Sage, London