

How to conduct a survey

Kevin Boone

School of Computing Science, Middlesex University

INDEX

1. Introduction 3

2. Questions and questionnaires 4

3. Sampling and the population 6

4. Certainty and confidence 7

5. Bias 8

6. Presentation of results 9

7. General guidelines 10

1. Introduction

Conducting a survey is often a useful way of finding something out, especially when 'human factors' are under investigation. Although surveys often investigate subjective issues, a well-designed survey should produce *quantitative*, rather than *qualitative*, results. That is, the results should be expressed numerically, and be capable of rigorous analysis.

This article describes some of the issues that the experimenter must confront when designing a small scale survey. By 'small scale' I mean one that can be carried out by a single person in a few weeks. Students very often underestimate how difficult is to carry out a survey well; a good survey is more than a handful of questionnaires and a couple of bar charts: it requires careful planning, methodical application, and detailed analysis of the results. In most surveys some statistical analysis will be required.

Perhaps this article should have been called 'how to start thinking about conducting a survey'. There is no general procedure that can be followed and which will automatically result in a good survey; all this article sets out to do is to bring the potential problems to the attention of the student. When planning a survey, students should find solutions to these problems which are compatible with the particular subject of the study, perhaps in consultation with their supervisors.

To illustrate the main points, I will be describing a hypothetical survey to find which of two different word processors is the easier to use. I will call these imaginary products *WordPro* and *WordPerfect*. Of course, these have no relation to the real products with the same names. Such surveys may well be conducted by companies that make word processors, in an attempt to improve the competitiveness of their products.

The most important issue to keep in mind when planning a survey is that you are trying to *find something out*. If you don't know in advance what the survey's objectives are, then you should question whether you really need the survey. The objectives of a survey can usually be phrased in the form of questions: 'Which word processor do most people use?' 'How long does it take to learn to use WordPro?' 'Why do people prefer one product over the other?' On the whole questions that start with 'Why...?' tend to be harder to answer than those that start with 'Which...?' or 'What...?'. They usually have to be translated into a series of 'What?' and 'How?' questions to be capable of rigorous interpretation.

2. Questions and questionnaires

If I wanted to find which word processor was easiest to use, I could simply ask people which they found easiest to use. In this case I would simply be passing the question along to the people in my survey group. If the questions are simple enough to answer that way, that's fine. Note, however, that 'Which word processor do you find easiest to use?' is a slightly different question from 'Which word processor is easiest to use?' To answer this second question I would have to find some *objective measure* of 'ease of use'. This measure might be based on measurements of time taken to carry out certain tasks, or number of documents produced in a given time, or whatever. Alternatively, I could frame a further set of questions that tries to get deeper insight into the issue of ease of use.

The important point here is that you will only get answers to the questions you ask. And even these will be contaminated with sampling error and bias, as we shall see. If the questions you ask do not satisfy the objectives of the survey, then the survey has failed.

To ask questions of a large number of people, many experimenters make use of questionnaires. In most cases, only a small number of people surveyed will respond, and the more complex the questionnaire the fewer responses there will be. The design of questionnaires is an issue about which complete books have been written. Here are a few general guidelines.

- If you are personally able to supervise the filling-in of each questionnaire, your survey group is probably inadequate. It is unlikely that 'family and friends of the experimenter' will be a representative sample of anything.
- If you are expecting people to reply by post, you will need to enclose a self-addressed envelope. Otherwise the recipient will almost certainly throw the questionnaire straight in the bin.
- If you can collect adequate results anonymously, then you should. In this case you should make it clear to potential respondents that anonymity will be assured. Even if you don't explicitly ask for a person's name, people will be reassured by the statement of confidentiality.
- If you can get meaningful results even if some questions are not answered, make this clear. A busy person is more likely to answer part of the questionnaire than all of it.
- You *must* ask the respondent for enough personal information to check that your sample group is adequately composed (see below). If you are asking a person to reply on behalf of an organization, then you need to ask these questions about the organization. Alternatively, you can mark or number the forms in some way that you can tell who has responded. However, doing so may deter potential responders.
- It is very easy to ask questions that are self-biasing. A trivial example is a question like 'do you object to filling in questionnaires?' If everyone replies 'no' you can't

claim that this generalizes: there could be any number of potential `yes' replies in people's waste paper baskets. While it is unlikely that anyone would make such an obvious mistake, more subtle examples exist to trap the unwary. For example, if you ask a group of people how much time they spend playing computer games at work, you will probably obtain a larger average than really exists. Why? Because people that have time to play computer games at work are more likely to find time to fill in your questionnaire. Here is another example. Some people like to use the World-Wide Web for providing questionnaires. The survey group will inevitably contain a larger-than-average proportion of computer and Internet enthusiasts, and people who are relatively knowledgeable about such matters.

There are no general solutions to this problem; you just have to think very carefully about your questions.

- People will not find time to fill in your questionnaire if they think it is a waste of time. People are more likely to think this if the questionnaire is poorly presented or ungrammatical. If your grammar and spelling are weak, get the questionnaire checked by someone else before you send it off.
- If you ask questions that have a bearing on a company's financial position, even trivial ones, you should not trust the replies. A person who is prepared to answer the questions is unlikely to have accurate information to answer them with. The people who are really in a position to answer such questions will not want to, as they will feel that this information could compromise the company's trading position. What's more, if there is even one question of this type, it will probably result in the whole questionnaire being discarded.

3. Sampling and the population

In almost all cases we would like to be able to *generalize* the results of a survey, that is, to estimate how the results might apply outside the survey group. Many students forget that, on the whole, the responses of the survey group itself are of *no interest whatever*. To know that a small group of people found one piece of software easier to use than another is not in itself interesting. What is interesting is to extend this finding to, say, computer users in general. We call this larger group the *population*. It isn't necessarily the same as the population of a country or the world, it simply means the group of people to whom the survey results should be extendable. In a survey of word processor preference, the population may be 'all users of word processors', or 'all people who use word processors at work', or something else. The important issue is that *you* have to decide, and plan the survey on the basis of your decision.

Beginners often forget that results of a small survey do not automatically extend to the population. There are two main reasons why this is so. First, there is *sampling variation*. Let's suppose that in our hypothetical survey 50 users of word processors were questioned; 60% of these people used WordPro and 30% used WordPerfect. The remaining 10% used something else entirely. It is likely that if a different group of 50 people were surveyed there would be different results, for example 55%/30%/15%. Both the survey groups gives results which are only *estimates* of the 'real', population-wide usage of word processors. Because they are only estimates, the two surveys are likely to give different results. This would be true even if the second group was alike in every measurable respect to the first. This is an inescapable principle of sampling: the survey group is only a *sample* of the general population, and measurements made on it are only estimates of the 'true' population values.

Second, in practice the second group surveyed will *not* be identical in all respects to the first. The group may consist of people of different ages, with different proportions of men and women, with different occupations, and so on. Any of these factors could affect the group's usage of word processors. How will this affect the result when the conclusions of the survey are extended to the general population? The result will be more accurate if the composition of the sample group is the same in all important respects to the composition of the population. If the sample group is very different to the population, we say it is *non-representative*. Non-representative sampling is one of the most frequent causes of error in surveys. For example, if I carry out a survey of votes cast at a polling station at eight o'clock in the morning, I will almost certainly get a different result compared with what I would have concluded at, say, ten o'clock in the morning. The reason is that fewer well-paid professionals are out and about at eight o'clock than manual workers; these groups of people will probably have different political viewpoints.

If you carry out a survey by selecting people whom it's convenient to question, then you have to accept that the results of the survey will only apply to the population of which this group is representative. For example, in the case of the 'word processor' survey, if I only question university students then my results will only be generalizable to other university students. Students that use word processors are not representative of 'people that use word processors'. The most obvious difference is that of age, although educational background will be an important difference as well. Often you can't do anything about this; but at the least you should recognize that the problem exists.

4. Certainty and confidence

So how many people does one need to survey to be *certain* that the survey results will apply to the whole population? The answer is that for certainty, we must survey the entire population. It's as simple as that. Of course in practice we usually can't do this, because there isn't enough time or money for such an undertaking. If the population in our hypothetical survey was 'all users of word processors', it could number several million individuals.

Because we can't have *certainty* we have to settle for *confidence*. The more people we survey, the more confident we become that the results apply to the population. A typical target is that of 95% confidence. Expressed simply, this means that we survey enough people that we can be 95% sure that the outcome applies to the population as well as the survey group. Note that '95% sure' in this case has a precise mathematical meaning. It means that if we repeated the survey many times, 19 times out of 20 ($=0.95$) we would obtain a result that was compatible with that of the population. If, for example it was true of the whole population that it found WordPro easier to use than WordPerfect, then a survey that had been chosen to give 95% confidence would obtain this correct result 19 times out of 20.

You can estimate the confidence level after carrying out the survey, but you must decide in advance what level of confidence will be acceptable. If your survey does not give this level of confidence you can use the results to plan a new, larger survey. If you cannot state what confidence you have in your results (as a number, not a vague hunch) then your survey is *worthless*.

Estimating confidence levels from a given set of data is a standard statistical procedure, and one that is described in any basic textbook on statistics. Estimating the size of the survey that will be needed to give the required confidence level is much more difficult, and requires consideration of the *statistical power* of your survey. Statistical power is a measure of how sensitive the survey result is to variations in the population, and is explained in slightly more advanced textbooks. Better still, consult a statistician. In any case, you need *some* data to estimate statistical power, and in most student projects this data will not be available at the outset.

Because it is difficult to estimate in advance how many people you need to survey, in many large projects there will be a *pilot study*, whose purpose is to find out enough about the population to plan the survey properly. In a student project you will probably have to settle for estimating the confidence levels at the end of the survey; if they are low you will need to find out why and suggest how they could be improved. Confidence levels are nearly always improved by increasing the size of the survey, but often a change in the survey design can give an improved confidence with much less expense.

5. Bias

A survey is *biased* if its outcome has been influenced by factors other than the one being studied. Bias is occasionally overt: the experimenter is not open-minded about the results, and interprets them wrongly. But more often bias comes from poor survey design. A typical problem is that of comparing two groups of people that are not really alike. For example, if there are more men than women in one group, and more women than men in another, the responses of the groups to any question will be influenced by the differences between men and women. In many cases these gender differences overwhelm the real subject of the study. Similar problems apply when groups have different age profiles.

The textbook solution to the problem of bias is that of randomization. This means picking survey subjects from the population group at random. Bear in mind that if you send out questionnaires and you use all the replies, this is not a random sample of anything. This is because people who take the trouble to respond to the questionnaire are probably not representative of the group you sent them to. In some cases it is necessary to use 'stratified' random sampling to ensure that the sample is typical of the population. For example, if I were surveying users of word processors, and I knew in advance that 60% of word processor users are women, I might want to ensure that 60% of people in my sample group were women. However, I would still try to select the individuals themselves at random.

6. Presentation of results

When presenting the results of a survey, you should try to include the minimum amount of data that communicates the overall findings effectively. If you are using questionnaires it is not usually helpful to include copies of every response. A summary of the responses is probably enough.

A problem that afflicts many students is that of distinguishing between 'objective' and 'subjective' reporting. It is quite important that a person reading the outcome of your survey can distinguish easily between factual or numerical results, and the experimenter's *interpretation* of the results. It is perfectly acceptable to conjecture about the reasons for a particular finding, but it is almost never helpful to mix facts and conjecture in a survey report. Bear in mind that the reader is also capable of interpreting your results, perhaps in a different way to you; to do this it needs to be easy to separate the objective results from your subjective interpretation.

The 'traditional' model for an experimental report has a section titled 'results' and one titled 'discussion'. The first of these is for plain, factual results and the second for interpretation and conjecture. This is still a sound way to report on the results of a survey.

If you use statistical analysis of your results, you don't need to include calculations, but you do need to include an explanation of the reason for adopting a particular statistical approach.

7. General guidelines

Here are some general guidelines that summarize the issues discussed in this article.

- Decide what the objectives of the study are. Ensure that you can phrase these objectives as questions or measurements. If you can't, *don't start*.
- Decide what to ask in your survey, or what to measure, that will give you answers to the questions you have selected. If the questions prove to be unanswerable, don't start.
- Ensure that you know what population you are studying. Pick a sample group that is representative of the population. If you cannot do this, at least try to estimate the effect that this lack of representativeness will have on your final results.
- If you can, pick a sample group large enough to give a result with a desired level of confidence or accuracy. If you can't do this, estimate the confidence or accuracy after the survey. If these are lower than you anticipated, make sure you can explain why.
- When presenting your results, be careful to distinguish between the factual or numerical results, and your interpretations of the results.