

Quantitative Methods in Social Science

Stephen Gorard

 **continuum**
NEW YORK • LONDON

Introduction: the role of numbers in research

WHY WE ALL NEED NUMBERS

A local paper recently ran a front-page story claiming that Cardiff was the worst area in Wales for unpaid television licences – it had ‘topped the league of shame for the second year running’. The evidence for this proposition was that there were more people in Cardiff (4,400) caught using TV without a licence than in any other ‘area’ of Wales (and it is important for readers to know that Cardiff is the largest city in Wales). Not surprisingly, the next worst area in the league of shame was Swansea (the second city of Wales), followed by Newport, then Wrexham, and so on. Everyone to whom I have told this story laughs at the absurdity of the claim and points out that the claim would have to be proportionate to the population of each area. Cardiff may then still be the worst, but at present we would have to assume that, as the most populous unitary authority in Wales, Cardiff would tend to have the most of *any* raw-score indicator (including, presumably, the number of people using TV *with* a licence). Why does this matter? It matters because very similar propositions are made routinely in social science research, and rather than being sifted out in peer review, they are publicized and often feted (see Chapter Three for some examples). This is indicative of the rather poor state of research involving very basic numbers – not that work like this gets done but rather that no one seems to care about the inconsistencies between the evidence and the conclusions drawn from it.

I have encountered books on all forms of social science research, some on statistical analysis and some on specialist topics such as survey design or sampling. There is not, to my knowledge, another practical book of advice for students on carrying out a research

project using quantitative techniques that links the three main methods of data derivation (secondary, survey and interventions) with their common methods of analysis. This is an important point, since the somewhat artificial separation of design and analysis leads to many of the common problems actually faced by students and those who deal with them (such as 'I have collected all this data, now please tell me what to do with it'). These issues are becoming more important as the climate in publicly funded research changes in favour of evidence-based policy and practice, with a growing interest in large-scale experimental trials and in the more general use of official data already collected for another purpose. This use of secondary data allows all students, perhaps for the first time, to carry out significant projects within a realistic timescale.

Above all, there is no book that steers a middle path of suggesting that *all* researchers should use numbers routinely in their research (even if only as 'consumers' of the quantitative research of others), while also cautioning against the potential artificiality of quantitative approaches and other associated perils. As well as laying out specific designs for both large- and small-scale social science research involving numbers, the book therefore also seeks to combat two idealized 'villains' – the student who does not 'do numbers' and is therefore forced to ignore all numeric results, and the student who is prepared only to 'do numbers' and tends to accept all numeric results at face value. Both extremes are common, in my experience, and dangerous. The emphasis throughout this book is therefore on selecting and using appropriate techniques, while considering the limitations inherent in any one approach. My underlying assumption is that there is no best method for social science research. There is simply differential fitness for purpose dependent upon the research question(s).

Some people have suggested that there should be more statistical ('quantitative') studies in social science research because this form of evidence is intrinsically preferable and of higher quality than other forms. I feel that this is completely the wrong way of looking at it. On the contrary, one reason to encourage a greater awareness of statistical techniques among all researchers is that quantitative work is currently often very poor, but largely unchecked. There are many other reasons why all researchers should learn something about techniques for research involving numbers. These reasons are outlined here and then presented in more detail throughout the book.

- So we won't get fooled again

The first and most obvious point is that the process of research involves some consideration of previous work in the same field. All researchers read and use the research of others. Therefore they need to develop what Brown and Dowling (1998) refer to as a 'mode of interrogation' for reading and using research results. If they do not have any understanding of research techniques involving numbers then they must either accept all such results without question, a very dangerous decision, or ignore all such results, a very foolish decision. In practice, many commentators attempt to create a middle way of accepting some results and rejecting others, even though they do not understand how the results were derived. This usually means that results are accepted on the basis of ideology or of whether they agree with what the commentator wants to believe. This is both dangerous *and* foolish. Whatever the people who do this like to call themselves, this is not a social scientific approach to research.

- Context is everything

Whatever your choice of primary method, there is a good chance that your research should involve numbers, at least at the outset. You may wish, for example, to document the experiences of the growing number of homeless people from ethnic minority backgrounds. Whatever approach you intend to use (participant observation, focus groups, anthropology, and so on) you should start from a quantitative basis. In order to direct your search you would use as much information as is available to you from the outset. You need to establish not only how many homeless people there are, but also where they are, how the socio-economic and ethnic patterns of these groups change over time and space, and so on. Such figures, termed 'secondary data', already exist, and therefore a preliminary analysis of them is the best place to start any study. Only then can you sensibly select a smaller group (a sample) for more detailed study. Existing figures, whatever their limitations, provide a context for any new study that is as important as the 'literature review' and the 'theoretical background'.

- Some techniques are common to all research

The use of a sample, for example, is a common phenomenon in all kinds of research using many different approaches to data collection and analysis. This book describes the process of sampling as it applies to all research involving samples, and is not specific to what have traditionally been considered as quantitative designs.

- We need an ideal

It is made clear in this book that experimental approaches have severe limitations in social science research. Nevertheless the ideal experiment, by isolating cause and effect, can provide us with a universal template for the perfect piece of research that leads to safe knowledge. We can then judge our more limited studies against that ideal, and so understand and explain the ways in which our own findings are less than secure (for sadly such is the fate of all real world research). True experiments may be rare in much social science research, but for the above reasons all researchers should still be able to design one (at least as a thought experiment). Even where an experiment is not used, we can adapt the formal logic of this scientific approach to deal with essentially passive approaches like observation (Boudon 1974). Once a discipline or field, like social science, is mature enough then some of its arguments can be converted into formal structures involving numbers. This helps to reduce ambiguity, clarify reasoning and reveal errors (see Chapter Seven).

- Because it is easy

Above all, it is important to realize that what is termed 'quantitative' research is generally very easy. Much analysis in social science involves nothing more complex than addition or multiplication – primary-school arithmetic in fact. Even this, along with any more complex calculations, is conducted for you by a computer. You have no need for paper and pencil. There is no need to practise any sums or memorize anything. Not only does this book *not* generally explain how to derive the formulae we use, generally it does not even state what those formulae are. These formulae are finished and complete. Therefore, no mathematics is involved in basic quantitative work. You can use statistics perfectly safely, just as you would drive a car without knowing or even caring how it works. There are always other books, software and expert advisers available to help if you 'break down'. The purpose of this book is to help explain when and how to use numeric techniques and how to report their results. The difficult bit lies in explaining your results and transforming them into practical reports for the users of research. This stage is, of course, common to all forms of research.

THE PENDULUM SWINGS

In 1988 *The Guardian* newspaper published an article called 'Who needs sociologists?', which described the near demise of the

discipline, and called for higher quality, less politically biased, and more relevant research. This led Marshall (1990) to comment that 'sociology ... is widely ridiculed by the ignorant ... and is regularly caricatured as left-wing rhetoric masquerading as scholarship'. To some extent, the latter position reflected the findings of the Rothschild (1982) report into the future of funding for UK social science research, which expressed 'disappointment' at progress in the field, and it also reflected the 'crisis of confidence' in all social sciences caused by the concurrent attacks of Sir Keith Joseph (then minister for Education and Science). These were linked to significant cuts in public funding for social science and even the threat of no funding at all, and were matched in other developed countries (Flather 1987). It was at this point that the Social Science Funding Council became the Economic and Social Research Council – removing the word 'science' from the title, perhaps as a sign of the political disdain for the soft methodologies of sociology in particular. Sociology is still not held in high general esteem, but perhaps the feeling is that little needs to be done about it because, unlike other fields, it seems to have little practical value. 'It is one thing having junk departments turning out junk sociologists, but quite another to be turning out junk engineers. If you think this is a point of no importance, imagine the next time you enter a lift...' (Brignell 2000, p. 12).

Over the last decade, the value and effectiveness of many other areas of social science research have been increasingly called into question (e.g. Lewis 2001, Hargreaves 1997, Tooley and Darby 1998). Educational research, for example, has been accused of being both 'second rate' and irrelevant to the needs and interests of practitioners. The Chief Executive of the Teacher Training Agency argued that 'despite the expenditure of over £65 million of public funding on educational research each year, there are surprisingly few studies which, individually or collectively, contribute systematically to the development of a comprehensive body of high quality evidence about pedagogy' (Millett 1997, p. 2). Research has been accused of being both 'second rate' and irrelevant to the needs and interests of practitioners. Her Majesty's Chief Inspector for Schools claimed to have given up reading research as 'life is too short. There is too much to do in the real world with real teachers in real schools to worry about methodological quarrels or to waste time decoding unintelligible, jargon-ridden prose to reach (if one is lucky) a conclusion that is often so transparently partisan as to be worthless' (Woodhead 1998, p. 51). This crisis of confidence is not confined to

the UK, having been pre-dated in the USA for example (Berliner and Biddle 1995, NRC 1999, NERPP 2000, Resnick 2000), nor to public policy research alone (Pirrie 2001, see also the fierce debates in anthropology, Tierney 2000). Indeed, it is currently characteristic of the relationship between the majority of professions and research, and there have been similar comments about the conduct of research in many public services (Dean 2000). Put simply, it seems that 'too many ... researchers produce second-rate work, and there are, for the most part, too few checks against this occurring' (Evans 2002, p. 44).

Of course, despite their public appeal, the evidence base for these criticisms is often weak, and this is part of what Marshall (1990) was writing about. However, these criticisms are general and strident enough for us to have to examine the quality of social science research. Part of the problem is an apparent system-wide shortage in expertise in large-scale studies, especially field trials derived from laboratory experimental designs. Over the last twenty years, there has undoubtedly been a move towards much greater use of 'qualitative' approaches (Hayes 1992), even in traditionally numerate areas of research (Ellmore and Woehilke 1998). In addition, acceptance rates for 'qualitative' publications are higher than for 'quantitative' pieces, by a ratio of around two to one in one US journal (Taylor 2001). There is a danger therefore of applying different standards of rigour to studies depending on their method and, presumably, on their referees. In some fields, the 1990s were dominated by generally small-scale funding leading to predominantly qualitative thinking (McIntyre and McIntyre 2000), entailing a considerable potential for bias (Dyson and Desforges 2002).

However, quantitative work has not stood still, and in the same period techniques for multivariate analysis, especially of data based on categories, have become considerably more sophisticated. While welcome, these twin developments may have increased the tendency towards a methodological schism, because individual researchers tend to specialize in one approach or the other. It is not unusual for one researcher never to have conducted any form of textual analysis and for another to admit to not having the least idea what 'multi-level modelling' is about, for example. Funders, such as the Economic and Social Research Council, of which Marshall is (at the time of writing) Chief Executive, want to see the pendulum swing back towards a more balanced portfolio of skills (e.g. Sooben 2002), and the ESRC currently has no fewer than fourteen initiatives in place to increase the use of quantitative approaches among social

scientists. Similar sentiments have been expressed in other developed countries (e.g. Diamond 2002). Part of the purpose of this book is to assist that swing.

INTRODUCING TWO VILLAINS

I have written this book as a general introduction to research design and statistical analysis for all students of social sciences. However, in doing so I have been particularly concerned to hinder the creation of two 'villainous' identities, both of which I meet regularly among students and even among more established researchers. They represent, if you like, two extreme viewpoints about numeric data – 'numbers are fab' and 'numbers are rubbish'.

Numbers are fab

This villain is perhaps most common in relatively established disciplines such as psychology, where there has been a tradition that only numeric data is of relevance. Students are therefore, perhaps unwittingly, encouraged to count or measure everything, even where this is not necessarily appropriate (as with some attitude scales, for example). One outcome is that statistical analysis is done badly and so gets a bad press. Allied to this approach is a cultural phenomenon I have observed, particularly with some international students and their sponsors, which again approves only research involving numbers. A corollary for both groups appears to be that forms of evidence not based on numbers are despised, while evidence based on numbers is accepted somewhat uncritically.

This last is clearly a problem, as I quite regularly come across findings that when reanalysed show the opposite to what is being claimed (e.g. Gorard 1997a, 2000a). In fact, I suspect that social science journals, books and edited chapters are full of quite basic arithmetic errors (and some of these are used for illustration throughout this book). Part of the problem here may be the 'cronyism' among reviewers that in-depth knowledge of advanced statistical procedures tends to generate, which leads to poorly explained and over-technical reports (where incomprehensible software-generated variable names are used routinely in descriptions of the analysis, for example).

As you will see throughout this book, I am a great fan of using computer software packages for statistical analysis, but the increasing quality and availability of these has exacerbated the problems outlined above in two ways. Software allows more and

more complex statistical models to be built and used, so that in the end most consumers of research simply cannot, or would not wish to, comprehend them. Even those who work on such high-level models have trouble transforming their findings into a format that does their analysis justice but also makes any sense to practitioners and policy-makers (see Goldstein et al. 2000 on the difficulties of this). This means that the 'average' consumer of research has either to implicitly accept the findings or to reject them as incomprehensible. Linked to the greater use of computers is the shotgun or dredging approach to analysis in which multiple exploratory analyses are run with the same set of data (see Chapter Nine). As well as liberating us from the drudgery of multiple calculations the computer has therefore increased the frequency of the 'blind or mindless application of methods without regard to their suitability for the solution of the problem at hand, or even in the complete absence of a clearly formulated problem' (Pedhazur 1982, p. 3).

Normal statistical textbooks describe ideal procedures to follow, but several studies of actual behaviour have observed different common practices among researchers. 'Producing a statistic is a social enterprise' (Gephart 1988 p. 15), and the stages of selecting variables, making observations and coding the results take place in everyday settings where practical influences arise. The divergence between the ideal and the actual is probably growing because of the increased accessibility to statistical software packages and a tendency to see these as 'expert systems' rather than convenient calculators. Statistical packages are making decisions for us that we may not even be aware of (through default settings). The possible dangers of this are increased because statistics have an under-stated rhetoric of their own, able to persuade specific audiences of their objectivity (Firestone 1987). The average researcher may be easily fooled by large numbers, confused by probabilities, prone to the fallacy of *post hoc ergo propter hoc*, and, without expertise of their own, led (and perhaps misled) by authorities (Brighton 2000). Perhaps this helps to explain why so few academic disputes over figures and subsequent corrections by authors appear in the literature.

Numbers are rubbish

The other villain is perhaps more common in the sociological tradition. Having realized that numbers can be used erroneously, sometimes even unscrupulously, some researchers simply reject all numeric evidence and its use (displaying what Mortimore and

Sammons [1997] call 'crude anti-quantitative attitudes', p. 185). This is as ludicrous a position as its opposite. As Clegg (1992) points out, we know that people sometimes lie to us but we do not therefore reject all future conversation. Why should lying with numbers be any different? I suspect, through my contact with students, that the key issue with numbers is a kind of fear or lack of confidence. But lack of confidence *can* be seen as a reasonably helpful characteristic for a researcher. It is surely better than the unjustifiable over-certainty represented by the 'numbers are fab' villain.

If we reject numeric evidence and its associated concerns about validity, generalizability and so on as the basis for research, then we are left with primarily subjective judgements. The danger therefore for 'qualitative' research conducted in isolation from numeric approaches is that it could be used simply as a rhetorical basis for retaining an existing prejudice. Without a combination of approaches we are often left with no clear way of deciding between competing conclusions. My argument is therefore not just that numeric evidence forms the basis of good qualitative studies and can be used to test its findings (the middle way, see Gorard 1998a). I am not even convinced that the very distinction between the two forms of evidence is a useful one (see the next section).

COMMON PROBLEMS IN RESEARCH

In each section of the book I illustrate some of the points being made through a consideration of problems I have encountered in my own research, the research of others and my work with novice researchers. To start with, here are three classic situations that you may find yourself in once you start to research.

- Being imprisoned by a 'paradigm'
- Deciding on a method before a topic
- Now ... how do I analyse all this?

Being imprisoned by a 'paradigm'

The term 'paradigm' is often applied to approaches to social science research. To my mind, this is never justified. Whatever its original value as a description of the 'chauvinism' that tends to appear in 'normal science' and the resistance to change in light of new ideas

(Kuhn 1970), the term has now done more harm than good to several generations of novice researchers. Instead of using 'paradigm' to refer to a topic or field of research (such as traditional physics) that might undergo a radical shift (to quantum physics, for example), people now use it to refer to a whole approach to research including philosophy, values and method. Moreover, and ironically of course, people tend to use the term to defend themselves against the need to change. Students, quite wrongly, can quickly become imprisoned in a 'paradigm' or feel they have to engage in pointless paradigm wars. They learn (because they are taught) that if they use any numbers in their research then they must be positivist or realist in philosophy, and they must be hypothetico-deductive or traditional in style. No one ever explains why these things are associated (apart from contingently). Texts making these bold claims apparently have no idea what terms like 'positivist' actually mean – Comte, the archetypal positivist, was against the use of statistical information in his 'social physics', for example (see also Steele 2002). If, on the other hand, students disavow the use of numbers in research then they must be interpretivist, holistic and alternative, believing in multiple perspectives rather than truth, and so on (e.g. Clarke 1999). This is such a common misunderstanding of the difference between the nature of numeric and non-numeric evidence and of the nature of truth, that it would require another whole book to discuss (but see Chapter Eleven). The important thing for the present is to consider that numbers can be used quite properly by *all* researchers whatever other methods they use. 'Qualitative and quantitative evidence' refers to a false dualism (Frazer 1995) and one that as researchers we would be better off without. One practical reason would be that we could cease wasting time and energy in pointless debates about the virtues of one approach over the other. Let's not be imprisoned by other peoples' ideas, at least until we have learnt a lot more about research in general.

The supposed distinction between qualitative and quantitative evidence is essentially a distinction between the traditional methods for their analysis rather than between underlying philosophies, paradigms or methods of data collection. As Heraclitus has written, 'logic is universal even if most people behave differently' (for if logic were not universal we could not debate with each other, so making research pointless). To some extent all methods of social science research deal with qualities, even when the observed qualities are counted. Similarly, all methods of analysis use some form of number, such as 'tend, most, some, all, none, few', and so on. This is what the

patterns in qualitative analysis are based on (even where the claim is made that a case is 'unique' since uniqueness is, of course, a numeric description). Words can be counted and numbers can be descriptive. Patterns are, by definition, numbers, and the things that are numbered are qualities (Popkewitz 1984). In fact, I sometimes wonder how many writers use qualitative analysis precisely to avoid the criticism that would be aimed at a more formal and transparent analysis. Examples of numeric analyses disguised as qualitative research appear later in this book.

Deciding on a method before a topic

Students have been heard to exclaim before deciding on a topic and research questions that they intend to use 'qualitative' methods of data collection or analysis, or that they are committed to the idea of a questionnaire. Perhaps 'it comes as no particular surprise to discover that a scientist formulates problems in a way which requires for their solution just those techniques in which he himself is especially skilled' (Pedhazur 1982, p. 28), but to understand this temptation is not to condone it. You must decide on your research topic and the questions you are curious about first, and only then consider how best to answer them. Don't fit your proposed study to your favourite approach (a case of the cart pulling the horse), and then try to disguise this as a philosophical, rather than a methodological decision (see above). This is another reason why all researchers need some knowledge of all methods.

Now ... how do I analyse all this?

Anyone who has dealt with student/novice researchers will have encountered this problem. In my institution this is not as frequent as it was, but I still see a reasonable number of people per year (perhaps sent by their supervisors for advice) who say, 'I have conducted a survey. Now can you tell me what to do with the answers?'. This is usually clear evidence of poor design. The reason that this book has alternate chapters on design and analysis is to try and help you see the two phases of research as concurrent. You cannot possibly design a sensible research instrument without considering in some detail how you will analyse the data you set out to collect. Otherwise you will not know if you have asked the right questions or collected data in the right format. The apparently separate phases of reading, formulating research questions, design, collection of data, analysis and reporting are really concurrent and iterative.

As outlined above, this book combines a consideration of the design and analysis of social science research involving numeric data. There is very little epistemology here. For those interested, my principles of research, such as they are, are very similar to the five norms described by Hammersley (1995, p. 76). I particularly like the first, which is that 'the overriding concern of researchers is the truth of claims, not their political implications or practical consequences'. For more on the philosophy of social science see Chapter Seven. For more on the ethical issues involved in research see Chapter Eight. For more about research 'paradigms' see Chapter Eleven. For a simple, sometimes amusing discussion of issues to put you in the 'right' frame of mind to grapple with research, see Fairbairn and Winch (1996), Huff (1991) and Thouless (1974). For a more serious approach to the abuse of statistics read Reichmann (1961). If you feel the need for some reminders about simple calculations see Solomon and Winch (1994). For a good introduction to social science research read Gilbert (1997), to formal statistics Clegg (1992) or Fielding and Gilbert (2000), and for help on writing a dissertation see Preece (1994) and many many others.

2

Finding secondary data: the 'idle' researcher

SUMMARY: USING SECONDARY DATA

As mentioned in Chapter One, one of the main reasons why all researchers are likely to need numbers, irrespective of their primary method, is that many of the large datasets available as context information for any study are numeric. The use of secondary data to help create or identify an appropriate sample (perhaps via stratification), to describe the pattern or problem to be explored by other methods, or even as a method in its own right, is growing. This is a trend encouraged by the funding councils, which allows cumulation and helps prevent the waste of resources involved in attempting research to explain non-existent patterns or problems. Existing statistics, whatever their limitations, provide a context for any new study, which is as important as the 'literature review' and the 'theoretical background'.

Consider this. I am not involved in running our university library, have never been to Newcastle and do not work for the Department of Education and Science. Nevertheless, without leaving the desk in my office, I could assemble within thirty minutes:

- a breakdown of the number and type of books borrowed at Cardiff University by the country of origin of all students (and therefore decide, for example, whether students from the western Pacific Rim read more books on statistics per year than those from the USA);
- an analysis of car ownership among the population of Newcastle broken down by the floor level of their permanent residence (and decide, for example, whether those living above first-floor level are less likely to own cars);
- a consideration of the rates of unauthorized absence from school