Research Design, as Independent of Methods

Stephen Gorard

School of Education

University of Birmingham

s.gorard@bham.ac.uk

Objectives

Readers of this chapter should be in a better position to:

- Understand the process of research design
- Place their own and others work within a full cycle or programme of ongoing research
- Understand why good research almost always involves a mixture of evidence
- Defend themselves from those who want numbers and text to be enemies rather than allies
- Argue that good research is more ethical for society than poor research

Introduction

The term 'mixed methods' is generally used to refer to social science research that combines evidence (or techniques or approaches) deemed both 'quantitative' and 'qualitative' in one study (e.g., Johnson and Onwuegbuzie, 2004; Creswell & Plano Clark, 2007). However mixed methods work is described, the key element seems to be this combination of quantitative and qualitative work at some level. It also appears that social science researchers as a body, and commentators on mixed methods in particular, view quantitative research as involving numbers, measurement and often forms of analysis based on sampling theory. Qualitative research, on the other hand, is viewed as almost anything not involving numbers, commonly text and discourse, but also images, observations, recordings and, on rare occasions, smell and other

sensory data. Each type of research has a specialist vocabulary, and an underlying philosophy, purportedly making each type of research a paradigm incommensurable with the other. Mixed methods approaches are therefore seen as complex, difficult and innovative because they seek to combine both of the q-word paradigms in the same study or research programme.

I was not fully aware of these paradigms, and their attendant beliefs, like positivism and interpretivism, when I started my PhD study as an educational practitioner. In that early study of school choice, I naturally used a variety of methods and techniques from simple re-analysis of existing datasets, documents, and archives through complex modelling of a bespoke survey, to in-depth observations and interviews (Gorard, 1997a). This seems to me what any novice researcher would do naturally (unless contaminated by the nonsense peddled in mainstream methods resources). Doing so seemed to cause me no problems, other than the time and effort involved, and I felt that my conclusions were drawn logically from an unproblematic synthesis of the various forms of evidence. It was only once I was underway that I discovered that methods experts believed what I was doing was wrong, impossible, or at least revolutionary in some way. In fact, what I did was none of those things. It seemed quite normal for anyone who genuinely wanted to find out the answer to their research questions, and from that time I began to try and explain to these experts and authorities why (Gorard,1997b). In the 12 years since my PhD I have completed about 60 projects large and small, and had about 600 publications of all types. In nearly all cases, I have continued to mix the methods known to others as quantitative and qualitative, both in the 'new political arithmetic' style of my PhD and in a variety of different styles including Bayesian syntheses, complex interventions, and design studies (e.g., Gorard, Taylor, & Fitz, 2003; Selwyn, Gorard, & Furlong, 2006; Gorard, et al. 2007). I have also continued to write about methods, including why quantitative work is misunderstood both by its advocates and by its opponents (e.g. Gorard, 2006, 2010), how misuse of the term 'theory' by advocates of qualitative research has become a barrier to mixed methods (e.g., Gorard, 2004a, 2004b), the ethics of mixing methods (e.g., Gorard, 2002a; Gorard with Taylor, 2004), and most importantly about the underlying universal logic of all research (e.g., Gorard, 2002b, 2002c).

Yet postgraduate students and new researchers in the UK are still routinely (mis-)taught about the incommensurable paradigms, and how they must elect one or other approach. Subsequently, they may be told that they can mix methods (if that is not a contradiction), and perhaps even told that a mixed methods approach is a third paradigm that they can choose (a bit like a fashion accessory). But the damage has been done by then. These new researchers may now view research evidence as a dichotomy of numbers and everything that is not numbers, and will reason that even if they can mix the two the fact of mixing suggests separate ingredients in the first place. If they are hesitant to work with numbers, they will tend to select the qualitative paradigm, and so convert their prior weakness in handling an essential form of evidence into a pretend bulwark and eventually a basis for criticising those who *do* use numbers. Those less hesitant with numbers will tend to find that the system both forces them to become quantitative (because it is only by opposites that the paradigms can be protective bulwarks) and positively encourages them as well, since there is a widespread shortage of social scientists willing and able to work with numbers in UK education research. For example, I review papers for around 50 journals internationally and I rarely get papers to review in the areas I work in. The common denominator to what I am sent is numbers. Editors send me papers with numbers not because I ask for them but because, unlike the majority of people in my field, I am prepared to look at them. Thus, I become, in their minds, a quantitative expert even though I am nothing of the sort, and have done as many interviews, documentary, archival, video and other in-depth analyses as most qualitative experts.

I believe that there is a different way of presenting the logic of research, not involving this particular unhelpful binary, through consideration of design and the full cycle of research work. I illustrate such an approach in this chapter, first looking at the relationship between methods and design and then between methods and the cycle. The chapter continues with a consideration of the differences between the q-word approaches. It ends with a consideration of the possible implications, if the argument thus far has been accepted, for the conduct of research, its ethics, and the preparation of new researchers. Of course, to my mind, the law of parsimony should mean that it is not necessary for me to argue in favour of an overall logic to social science research with no schism and no paradigms (as that term is used here, rather than as the fluid conversion of questions into puzzles as discussed by Kuhn and

others). But it may be interesting for those imbued with 'isms' at least to understand my point of view.

Research design in social science: the forgotten element?

Research design in the social sciences, as elsewhere, is a way of organising a research project or programme from inception in order to maximise the likelihood of generating evidence that provides a warranted answer to the research questions for a given level of resource. The emphasis is less on how to conduct a type of research than on which type is appropriate in the circumstances (Hakim, 2000). In the same way that research questions can evolve as a project unfolds, so can its design(s). The structure of a standard design is not intended to be restrictive, since designs can be easily used in combination; nor is it assumed that any off-the-shelf existing design is always or ever appropriate. Instead, consideration of design at the outset is intended to stimulate early awareness of the pitfalls and opportunities that will present themselves, and through knowledge of prior designs to simplify subsequent analysis, and so aid warranted conclusions.

There are many elements to consider in a research design, but they commonly include the treatment or programme to be evaluated (if there is one), the data collected, the groups and sub-groups of interest, the allocation of cases to groups or to treatments (where appropriate), and what happens over time (unless the study is a snapshot). Any design or project may have only some of these elements. Perhaps the most common type of design in social science involves no treatment, no allocation to groups, and no consideration of time. It is cross-sectional with one or more pre-existing groups. It is also often described as 'qualitative' in the sense that no measurements are used, and the data are often based on interviews. It is this design that makes it hardest to warrant any claims, since there is usually no comparison between groups or over time or place. But actually this design does not entail any specific kind of data, any more than any other design. In fact, nothing about consideration of treatments, data collection, groups, allocation and time entails a specific kind of evidence. Most designs seem to me to be an encouragement to use a variety of data. Standard designs can be classified in a number of ways, such as

4

whether there is a treatment or intervention (active) or not (passive). Active designs include:

- Randomised control trials (with or without blinds)
- Quasi-experiments – including interrupted time series
- Natural experiments
- Action research
- Design studies

and some might say

- Participant observation.

Passive designs include:

- Cohort studies (time series and retrospective)
- Other longitudinal designs
- Case-control studies
- New political arithmetic
- Cross-sectional studies

and some might say

- Systematic reviews (including Bayesian)

The choice depends largely on the kind of claims and conclusions to be drawn, and to a lesser extent on the practicalities of the situation and resources available. I say these are lesser considerations because if it is not possible for financial, ethical or other reasons to use a suitable design then the research should not be done at all (as opposed to being done badly, and perhaps leading to inappropriate claims to knowledge). The need for warranted conclusions requires the researcher to identify the kind of claims to be made – such as descriptive, associative, correlational or causal – and then ensure that the most appropriate possible design is used. Put simply, a comparative claim must have an explicit and suitable comparator, for example. The warranting principle is based on this consideration - if the claim to be drawn from the evidence is not actually true then how else could the evidence be explained? The claim should be the simplest explanation for the available evidence.

What the design should do is eliminate (or at least test or allow for) the greatest possible number of alternate explanations. In this way, the design eases the analysis process, and provides part of the warrant for the research claims.

What all of these designs, and variants of them, have in common is that they do not specify the kind of data to be used or collected. At the level of an individual study, the research design used by social scientists will be independent of, and logically prior to, the methods of data collection and analysis employed. No kinds of data, and no particular philosophical predicates, are entailed by common existing design structures such as longitudinal, case study, randomised controlled trial, or action research. A good intervention, for example, could and should use a variety of data collection techniques to understand whether something works, how to improve it, or why it does not work. Experiments can use any kind of data as outcomes, and collect any kind of data throughout to help understand why the outcomes are as they are. Longitudinal studies can collect data of all types over time. Case studies involve immersion in one real-life scenario, collecting data of any kind ranging from existing records to *ad hoc* observations. And so on.
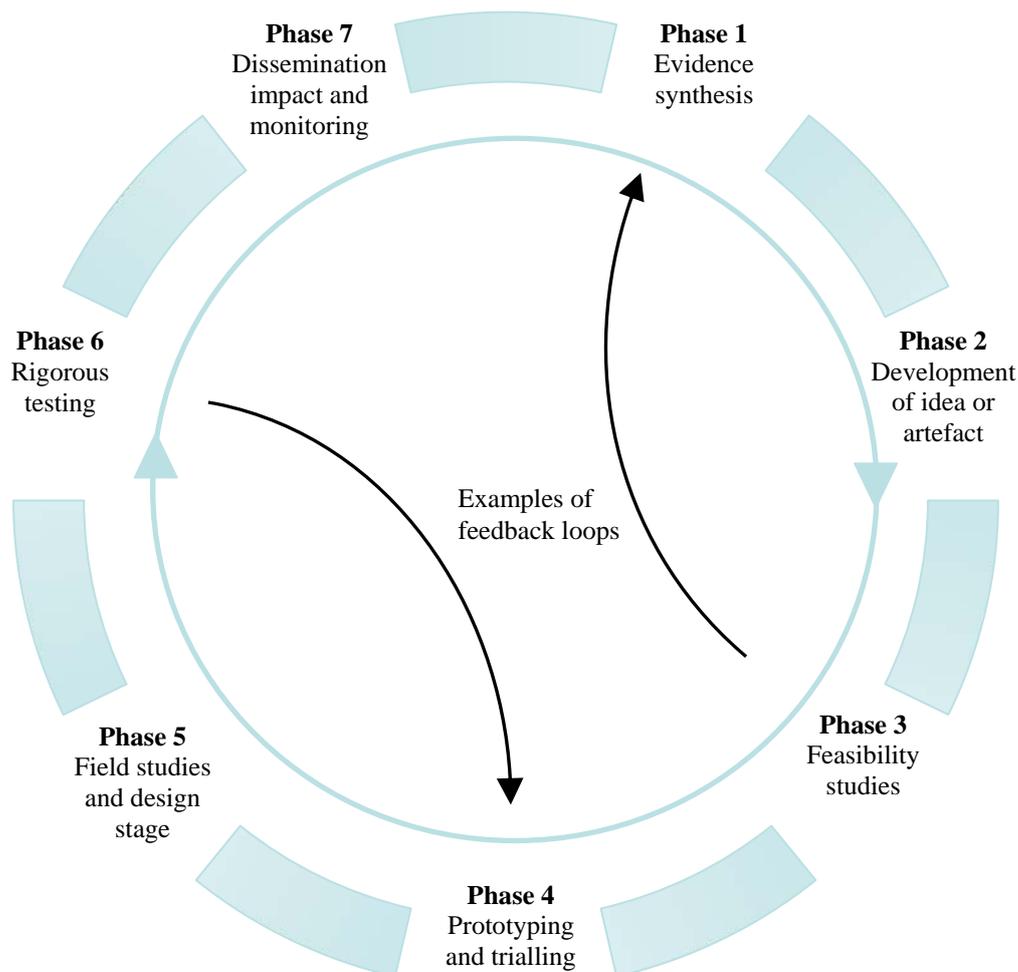
Mixed methods approaches are therefore not a kind of research design; nor do they entail or privilege a particular design. Of course, all stages in research can be said to involve elements of 'design'. The sample design is one example, and the design of instruments for data collection another. But research design, as usually defined in social science research, and as discussed here, is a prior stage to each of these (de Vaus, 2001).

The cycle of research

At the meta-level of a programme of research conducted by one team, or a field of research conducted by otherwise separate teams, the over-arching research design will incorporate most methods of data collection and analysis. Figure 1 is a simplified description of a full cycle for a research programme (for a fuller description and discussion, see Middleton, Gorard, Taylor, & Bannan- Ritland, 2008). It is based on a number of sources, including the genesis of a design study (Gorard with Taylor,

2004), the UK Medical Research Council model for undertaking complex medical interventions (MRC, 2000) and one OECD conception of what useful policy research looks like (Cook & Gorard, 2008). The cycle is more properly a spiral which has no clear beginning or end, in which activities (phases) overlap, can take place simultaneously, and iterate. Nevertheless, the various phases should be recognisable to anyone working in areas of applied social science, like public policy. Starting with draft research questions, the research cycle might start with a synthesis of existing evidence (phase 1 here). Ideally this synthesis would be an inclusive review of the literature both published and unpublished (perhaps combining the different kinds of evidence via a Bayesian approach – Gorard, Roberts, & Taylor, 2004), coupled with a re-analysis of relevant existing datasets of all kinds (including data archives and administrative datasets), and related policy/practice documents. It is not possible to conduct a fair appraisal of the existing evidence on almost any real topic in applied social science without naturally combining evidence involving text, numbers, pictures, and a variety of other data forms. Anyone who excludes relevant data because of its type (such as text or numeric) is a fake researcher, not really trying to find anything out.

Figure 1 – An outline of the full cycle of social science research and development

Phase 7
Dissemination
impact and
monitoring

Phase 1
Evidence
synthesis

Phase 6
Rigorous
testing

Phase 2
Development
of idea or
artefact

Examples of
feedback loops

Phase 5
Field studies
and design
stage

Phase 3
Feasibility
studies

Phase 4
Prototyping
and trialling

Currently, the kind of comprehensive synthesis outlined above is rare. If more took place, one consequence might be that a research programme more often ended at phase 1, where the answers to the research questions are already as well established as social science answers can be. Another consequence might be that researchers more often revised their initial questions suitably before continuing to other phases of the cycle (White, 2008). For example, there is little point in continuing to investigate why the attainment gap between boys and girls at school is increasing if initial work shows that the gap is actually decreasing (Gorard, Rees, & Salisbury 2001). The eclectic reuse of existing evidence would often be more ethical than standard practice (a patchy literature review), making better and more efficient use of the taxpayer and charitable money spent on the subsequent research.

Similarly, where a project or programme continues past phase 1, every further phase in the cycle would tend to involve a mixture of methods. Each phase might lead to a

realisation that little more can be learnt and that the study is over, or that the study needs radical revision and iteration to an earlier phase(s), or progression to a subsequent phase. The overall programme might be envisaged as tending towards an artefact or 'product' of some kind. This product might be a theory (if the desired outcome is simply knowledge), a proposed improvement for public policy, or a tool/resource for a practitioner. In order for any of these outcomes to be promoted and disseminated in an ethical manner they must have been tested (or else the dissemination must merely state that they seem a good idea but that we have no real idea of their value). A theory, by definition, will generate testable propositions. A proposed public policy intervention can be tested realistically and then monitored *in situ* for the predicted benefits, and for any unwanted and undesirable side effects. Therefore, for that minority of programmes which continue as far as phase 6 in Figure 1, rigorous testing must usually involve a mixture of methods and types of evidence in just the same way as phase 1. Even where a purely numeric outcome is envisaged as the benefit of the research programme (such as a more effective or cost-efficient service) it is no good knowing that the intervention works if we do not also know that it is unpopular and likely to be ignored or subverted in practice. Similarly, it would be a waste of resource, and therefore unethical, simply to discover that an intervention did not work in phase 6 and so return to a new programme of study in phase 1. We would want to know why it did not work, or perhaps how to improve it, and whether it was effective for some regular pattern of cases but not for others. So in phase 6, like phase 1, the researcher or team who genuinely wants to find something out will naturally use a range of methods and approaches including measurement, narrative and observation.

The same kind of conclusion could be reached for every phase in Figure 1. Even monitoring and evaluation of the rollout of the results (phase 7) is best done by using all and any data available. Even in simple academic impact terms, a citation count for a piece of research gives no idea of the way in which it is used (just mentioned or fundamental to the new work of others), nor indeed whether the citation is critical of the research and whether it is justified in being critical. On the other hand, for a widely cited piece of research, reading in-depth how the research has been cited in a few pieces gives no idea of the overall pattern. Analysing citation patterns *and* reading some of the citing pieces – perhaps chosen to represent features of the overall

pattern – gives a much better indication of the impact of this research. Methods of data collection and analysis are not alternatives; they are complementary. Specific methods might be used to answer a simple, perhaps descriptive, research question in one phase, but even then the answer will tend to yield more complex causal questions that require more attention to research design (Cook and Gorard, 2008).

Across all stages of the cycle up to definitive testing, engineering of results into useable form, and subsequent rollout and monitoring, different methods might have a more dominant role in any one stage, but the overall process for a field of endeavour requires a full range of research techniques. It is indefensible for a researcher, even one limited in expertise to one technique (and so admitting that they are not competent to conduct even something as basic as a comprehensive literature review, for example), to imagine that they are not involved in a larger process that 'mixes' methods naturally and automatically.

Re-considering the schism

Therefore, the q-word dichotomy has, as illustrated, no relevance to design or indeed to entire programmes of research. We may consider that surveys and interviews, for example, are quite different, but even here there may be a continuum through structured interview schedules to open-ended survey items delivered face-to-face. The q-word division is not helpful even with methods. Is there such a thing as a qualitative interview *and* a quantitative interview? I doubt it. Interview, as a general category, is enough. The q-words add nothing. So what lies beneath the schism? I consider here three general propositions – that the schism arises from important differences in paradigm, scale, and methods of data analysis.

The q-words are not paradigms

In the sociology of science the notion of a 'paradigm', is a description of the sets of socially accepted assumptions that tend to appear in 'normal science' (Kuhn, 1970). A paradigm is a set of accepted rules within any field for solving one or more puzzles – where a puzzle is defined as a scientific question that it is possible to find a solution to

in the near future, to distinguish it from the many important and interesting questions that do not have an answer at any particular stage of progress (Davis, 1994). 'Normal science' in Kuhnian terms is held together, rightly or wrongly, by the norms of reviewing and acceptance that work in that taken-for-granted theoretical framework. A paradigm shift occurs when that framework changes, perhaps through the accumulation of evidence, perhaps due to a genuinely new idea, but partly through a change in general acceptance. Often a new paradigm emerges because a procedure or set of rules has been created for converting another more general query *into* a puzzle. But, what Kuhn saw as normal science could also be simply passive and uncritical rather than genuinely cumulative in nature. It could be based on practices that differ from those stated, because of deceit, either of the self or the audience (Lakatos, 1978, p.44), and because researchers conceal their actual methodological divergence in practice (Gephart, 1988).

However, instead of using 'paradigm' to refer to a topic or field of research (such as traditional physics) which might undergo a radical shift on the basis of evidence (to quantum physics, for example), some commentators now use it to refer to a whole approach to research including philosophy, values and method (Perlesz and Lindsay, 2003). The most common of these approaches are qualitative and quantitative, even though the q-words only make sense, if they make sense at all, as descriptions of data. These commentators tend to use the term paradigm conservatively, to defend themselves against the need to change, or against contradictory evidence of a different nature to their own. Their idea of paradigm appears to defend them because they pointlessly parcel up unrelated ideas in methodology (as explained in Chapter Four of this collection). The idea of normal science as a collection of individuals all working towards the solution of a closely defined problem has all but disappeared. Instead, we have paradigm as a symptom of scientific immaturity. The concept of paradigm has, thus, become a cultural cliché with so many meanings it is now almost meaningless. And many of the terms associated with paradigms – the 'isms' such as positivism – are used almost entirely to refer to others, having become intellectually acceptable terms of abuse and ridicule (see also Hammersley, 2005).

Unfortunately, some novice research students can quickly become imprisoned within one of these fake qualitative and quantitative 'paradigms'. 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 (see, for example, such claims by Clarke, 1999). 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. Sale, Lohfeld, and Brazil (2002), for example, claim that 'The quantitative paradigm is based on positivism. Science is characterized by empirical research' (p.44). Whereas, 'In contrast, the qualitative paradigm is based on… multiple realities. [There is] no external referent by which to compare claims of truth' (p.45). Such commentators 'evidently believe that the choice of a research method represents commitment to a certain kind of truth and the concomitant rejection of other kinds of truth' (Snow, 2001, p.3). They consider that the value of their methods can be judged completely separately from the questions they are used to answer.

What is ironic about this use of the term 'paradigm' to refer to a methods- and value-based system in social research is that it has never been intended to be generally taken-for-granted, in the way that 'normal science' is. Rather, it splits the field into two non-communicating parts. Therefore, a paradigm of this kind cannot be shifted by evidence, ideas, or the fact that others reject it. It becomes divisive and conservative in nature, leading to 'an exaggeration of the differences between the two traditions' (Gray & Densten, 1998, p.419) and an impoverishment of the range of methods deployed to try and solve important social problems.

It is somewhat impractical to sustain an argument that all parts of all methods, including data collection, carry epistemological or ontological commitments anyway (Frazer, 1995; Bryman, 2001). So, researchers tend to confuse the issues, shuttling from technical to philosophical differences, and exaggerating them into a paradigm (Bryman, 1988). No research design implies either qualitative or quantitative data even though reviewers commonly make the mistake of assuming that they do – that experiments can only collect numeric data, observation must be non-numeric, and so on. Observation of how work is conducted shows that qualitative and quantitative work are *not* conducted in differing research paradigms, in practice. The alleged differences between research paradigms (in this sense) prevail in spite of good evidence, not because of it (Quack theories, 2002).

Mixed methods have been claimed to be a third paradigm (Johnson & Onwuegbuzie, 2004), but this seems to add to the confusion by apparently confirming the validity of the first two, instead of simply blowing them all away by not mentioning any of them in the development of new researchers. World views do not logically entail or privilege the use of specific methods (Guba, 1990), but may only be thought to be so due to a common confusion between the logic of designing a study and the method of collecting data (according to de Vaus, 2001; Geurts & Roosendaal, 2001). 'The researcher's fidelity to principles of inquiry is more important than allegiance to procedural mechanics... Research should be judged by the quality and soundness of its conception, implementation and description, not by the genre within which it is conducted' (Paul & Marfo, 2001, pp. 543-545). In real-life, methods *can* be separated from the epistemology from which they emerge, so that qualitative work is not tied to a constructivist paradigm, and so on (Teddlie & Tashakkori, 2003). The paradigm argument for the q-word approaches is a red herring, and unnecessarily complex to boot (as evidenced in some of the other chapters in part one of this collection).

Not just an issue of scale

Some authorities suggest that a clear difference between the q-word approaches is their scale (e.g., Creswell & Plano Clark, 2007), with qualitative data collection necessarily involving small numbers of cases, whereas quantitative relies on very large samples in order to increase power and reduce the standard error. This is misleading for two reasons. First, it is not an accurate description of what happens in practice. Both Gorard & Rees (2002) and Selwyn, Gorard & Furlong (2006) interviewed 1,100 adults in their own homes, for example, and treated the data gathered as individual life histories. This is larger-scale than many surveys. On the other hand, Smith & Gorard (2005) conducted a field trial in one school with only 26 students in the treatment group, yielding both attainment scores and contextual data. The number of cases is not necessarily related to methods of data collection or to either of the q-words. Second, issues such as sampling error and power only relate to a tiny minority of studies where a true and complete random sample is used or where a population is randomly allocated to treatment groups. In the much more common situations of working with incomplete samples with measurement error or dropout,

convenience, snowball and other non-random samples and the increasing amount of population data available to us, the constraints of sampling theory are completely irrelevant. It is also the case that the standard error/power theory of analysis is fatally flawed in its own terms, even when used as intended (Gorard, 2010). The accounts of hundreds of interviewees can be properly analysed as text, and the account of one case study can properly involve numbers. The supposed link between scale and paradigm is just an illusion.

The logic of analysis is similar

Another possible distinction between the q-word approaches is their method of analysis. Qualitative work is supposed to be subjective and so closer to a social world (Gergen & Gergen, 2000). Quantitative work is supposed to help us become objective (Bradley & Schaefer, 1998). This distinction between quantitative and qualitative analysis is exaggerated, largely because of widespread error by those who do handle numbers (Gorard, 2010) and ignorance of the subjective and interpretivist nature of numeric analysis by those who do not (Gorard, 2006). The similarities of the underlying procedures used are remarkable (Onwuegbuzie and Leech, 2005). Few analytical techniques are restricted by data gathering methods, input data, or by sample size. Most methods of analysis use some form of number, such as 'tend', 'most', 'some', 'all', 'none', 'few' and so on (Gorard, 1997b). Whenever one talks of things being 'rare', 'typical', 'great' or 'related' this is a numeric claim, and can only be so substantiated, whether expressed verbally or in figures (Meehl, 1998). Similarly, quantification does not consist of simply assigning numbers to things, but of relating empirical relations to numeric relations (Nash, 2002). The numbers themselves are only valuable insofar as their behaviour is an isomorph of the qualities they are summarising. Statistical analysis is misunderstood by observers if they do not consider also the social settings in which it takes place, and the role of 'qualitative' factors in reaching a conclusion (MacKenzie, 1999). 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 subjective influences arise. It would be dishonest to pretend otherwise.

Even such an apparently basic operation as the measurement of a length involves acceptance of a series of theories and judgements about the nature of length and the isomorphic behaviour of numbers (Berka, 1984). As with 'number' and 'length', so also with many of our basic concepts and classifications for use in social science – 'sex', 'time', 'place', 'family', 'class' or 'ethnicity' (Gorard, 2003). Measurement is an intrinsically intrepretivist process (Gorard, 2009). Personal judgement(s) lie at the heart of all research – in our choice of research questions, samples, questions to participants and methods of analysis – regardless of the kinds of data to be collected. The idea that the quantitative work is objective and qualitative is subjective is based on a misunderstanding of how research is actually conducted.

Implications (if the argument so far is accepted)

For the conduct of research

Mixed methods are not a design. Nor do they represent some kind of paradigm, separate from those traditionally termed 'qualitative' and 'quantitative'. How could mixed methods be incommensurable with the two elements supposed to be mixed within them? Mixed methods are then just a description of how most people would go about researching any topic that they really wanted to find out about. The results of research if taken seriously affect the lives of real people, and lead to genuine expenditure and opportunity costs. We should be (nearly) as concerned about research as we are about investigations and decisions in our lives. It is instructive to contrast how we, as researchers, generally behave when conducting research professionally and how we behave when trying to answer important questions in our everyday lives. When we make real-life decisions about where to live, where to work, the care and safety of our children, the health of our loved ones, and so on, many of us behave very differently from being 'researchers'.

No one, on buying a house, refuses to discuss or even know the price, the mortgage repayments, the room measurements or the number of bathrooms. No one, on buying a house, refuses to visit the house, look at pictures of it, walk or drive around the

neighbourhood, or talk to people about it. All rational actors putting a substantial personal investment in their own house would naturally and without any consideration of paradigms, epistemology, identity or mixed methods, use all and any convenient data to help make up their mind. We will believe that the house is real even though external to us, and that it remains the same even when we approach it from different ends of the street. Thus, we would not start with 'isms'. We would not refuse to visit the house, or talk to the neighbours about it, because we were 'quantitative' researchers and did not believe that observation or narratives were valid or reliable enough for our purposes. We would not refuse to consider the interest rate for the loan, or the size of the monthly repayments, because we were 'qualitative' researchers and did not believe that numbers could do justice to the social world. And we would naturally, even unconsciously, synthesise the various forms of data to reach a verdict. I do not mean to say that such real-life decisions are easy, but that the difficulties do not stem from paradigms and epistemology, but from weighing up factors like cost, convenience, luxury, safety etc. People would use the same naturally mixed approach when making arrangements for the safety of their children or loved ones, and for any information-based task about which they really cared. For important matters, we behave sensibly, eclectically, critically, sceptically, but always with that final leap of faith because research, however carefully conducted, does not provide the action - it only informs the action. We collect all and any evidence available to us as time and resources allow, and then synthesise it naturally, without consideration of mixing methods as such.

Thus, I can envisage only two situations in which a social science researcher would not similarly use 'mixed methods' in their work. Perhaps they do not care about the results, and are simply pretending to do research (and wasting peoples' time and money in the process). This may be a common phenomenon in reality. Or their research question is peculiarly specific, entailing only one method. However, the existence of this second situation, analogous to using only one tool from a larger toolbox, is not any kind of argument for separate paradigms of the two q-words and mixed methods. Mixed methods, in the sense of having a variety of tools in the toolbox and using them as appropriate, is the only sensible way to approach research. Thus, a central premise of mixed methods is that 'the use of quantitative and qualitative approaches in combination provides a better understanding of research

problems than either approach alone' (Creswell & Plano Clark, 2007, p.5). This is what I have always argued, but without the need to create a new paradigm (Gorard, 1997a; Gorard with Taylor, 2004). Mixed methods (the ability to use any appropriate methods) is the only sensible and ethical way to conduct research.

For ethical consideration of projects

A key ethical concern for those conducting or using publicly-funded research ought to be the quality of the research, and so the robustness of the findings, and the security of the conclusions drawn. Until recently, very little of the writing on the ethics of education research has been concerned with quality. The concern has been largely for the participants in the research process, which is perfectly proper, but this emphasis may have blinded researchers to their responsibility to those not participating in the research process. The tax-payers and charity-givers who fund the research, and the general public who use the resulting public services, for example, have the right to expect that the research is conducted in such a way that it is possible for the researcher to test and answer the questions asked The general public, when this is demonstrated, are shocked to discover that they are funding the work of social scientists who either believe that everything can be encompassed in numbers, or much more often who believe that nothing can be achieved using numbers (or that nothing is true, or that there is no external world, or….).

Generating secure findings for widespread use in public policy should involve a variety of factors including care and attention, sceptical consideration of plausible alternatives, independent replication, transparent prior criteria for success and failure, use of multiple complementary methods, and explicit rigorous testing of tentative explanations. The q-word paradigms are just a hindrance here, and so are unethical as originally suggested in (Gorard, 2002a, 2003), with this second principle of research ethics slowly filtering into professional guidelines (e.g. Social Research Association 2003).

For the development of new researchers

As I explained at the start of the chapter, I was lucky enough to be undamaged by supposed research methods development of the kind now compulsory for publicly-funded new researchers in the UK. Or perhaps I was critical and confident enough to query what methods experts were saying and writing. Methods text, courses and resources are replete with errors and misinformation, such that many do more damage than good. Some mistakes are relatively trivial. I remember clearly being told by international experts that triangulation was based on having three points of view, or that the finite population correction meant that a sample must be smaller than proposed, for example. I have heard colleagues co-teaching in my own modules tell our students that regression is a *test* of causation (see also Robinson, et al. 2007), or that software like Nvivo will analyse textual data for them. Some examples are more serious. There is a widespread error in methods texts implicitly stating that the probability of a hypothesis given the data is the same as, or closely related to, the probability of the data given that the hypothesis is true. However, probably the most serious mistakes currently made in researcher development are the lack of awareness of design, and the suggestion that methods imply values, and are a matter of personal preference rather than a consequence of the problems to be overcome via research.

Much research methods training in social science is predicated on the notion that there are distinct categories of methods such as 'qualitative' or 'quantitative'. Methods are then generally taught to researchers in an isolated way, and this isolation is reinforced by sessions and resources on researcher identities, paradigms, and values. The schism between qualitative and quantitative work is very confusing for student researchers (Ercikan and Wolff-Michael, 2006). It is rightly confusing because it does not make sense. These artificial categories of data collection and analysis are not paradigms. Both kinds of methods involve subjective judgements about less than perfect evidence. Both involve consideration of quantity and of quality, of type and frequency. Nothing is gained by the schism, and I have been wrong in allowing publishers to use the q-words in the title of some of my books (altering 'The role of number made easy' to 'Quantitative methods', for example). Subsequently, many of the same methods training programmes taken by new researchers refer to the value of mixing methods, such as those deemed 'qualitative' or 'quantitative'. Perhaps, unsurprisingly, this leads to further confusion. Better to

leave paradigms, schisms and mixing methods for later, or even leave them out of training courses altogether.

It is not enough merely to eliminate the q-words from module headings and resources. The change has to be adopted by all tutors respecting every kind of evidence for what it is, and following this respect through in their own teaching and writing. This is what I have implemented successfully in both previous universities in which I have worked. It is what I am trying to implement in my current institution – encouraged as ever by national funding bodies, supported by the upper echelons of the university, and opposed by the least research-active of my colleagues who seem to want to cling their comforting paradigms, perhaps as an explanation for their unwillingness to conduct relevant, rigorous and ethical research. This is part of the reason why I would want research methods development for new researchers to be exclusively in the hands of the most successful practical researchers, who are often busy doing research, rather than in the hands of those supposed methods specialists, who are often unencumbered by research contracts and so free to corrupt the researchers of the future. Busy practical researchers will tend to focus on the craft, the fun, the importance *and* the humility of research. They will want new researchers to help them combat inequality, inefficiency, and corruption in important areas of public policy like health, education, crime and social housing. There is just no time to waste on meaningless complications and the cod philosophy of the q-word paradigms.

Summary

This chapter looks at the idea of mixed methods approaches to research and concludes that this is the way new researchers would naturally approach the solution of any important evidence-informed problem. This means that a lot of the epistemology and identity routinely taught to new researchers is not just pointless; it may actually harm their development. The chapter reminds readers of the importance of research design, and how this neglected stage of the research cycle is completely independent of issues like methods of data collection and analysis. The schismic classifications of qualitative and quantitative work are unjustifiable as paradigms.

They are not based on the scale of the work, nor on different underlying logic of analysis. They are pointless. The chapter ends with some considerations of the implications for the conduct of publicly-funded research, for the ethics of social science, and for the preparation of new researchers.

Related questions

1. If the first principle of research ethics is not to harm research participants, how would you summarise the second principle discussed in ths chapter?
2. Can all issues of research ethics be classified under these two principles, or are there more?
3. Why do you think so many professional researchers think it is possible to claim that researchers should ignore either evidence in the form of text or evidence in the form of numbers?
4. Try to imagine a real-life situation that is important to you in which you had to make an evidence-informed decision. What reason could you have for ignoring relevant evidence simply because it was numeric (or textual)?
5. Look at some journals in your area of interest and consider how many papers use techniques based on random sampling theory (such as significance tests, standard errors, confidence intervals). How many of these actually had random samples, and how many were using these techniques erroneously?
6. Look at some journals in your area of interest and consider how many papers using purportedly 'qualitative' methods make either explicit or implicit comparative claims (over time, place or social group) without presenting any data from a comparator group?
7. Examine the meaning and use of the term 'warrant' in social science research. How useful is it for your own work?

References

Berka, K. (1983). *Measurement: its concepts, theories and problems*, London: Reidel.

Bradley, W. and Shaefer, K. (1998). *Limitations of Measurement in the Social Sciences*, California: Sage.

Bryman, A. (1988). *Quantity and quality in social research*, London: Unwin Hyman.

Bryman, A. (2001). *Social research methods*, Oxford: Oxford University Press.

Clarke, A. (1999). *Evaluation research*, London: Sage.

Cook, T. and Gorard, S. (2007). What counts and what should count as evidence, pp.33-49 in OECD (Eds.). *Evidence in education: Linking research and policy*, Paris: OECD.

Creswell, J. and Plano Clark, V. (2007). *Designing and conducting mixed methods research*, London: Sage.

Davis, J. (1994). What's wrong with sociology?, *Sociological Forum*, 9, 2, 179-197.

de Vaus, D. (2001). *Research design in social science*, London: Sage.

Ercikan, K. and Wolff-Michael, R. (2006). What good is polarizing research into qualitative and quantitative?, *Educational Researcher*, 35, 5, 14-23.

Frazer, E. (1995). What's new in the philosophy of science?, *Oxford Review of Education*, 21, 3, 267.

Gephart, R. (1988). *Ethnostatistics: Qualitative foundations for quantitative research*, London: Sage.

Gergen, M. and Gergen, K. (2000). Qualitative Inquiry, Tensions and Transformations, in N. Denzin and Y. Lincoln (Eds.). *The Landscape of Qualitative Reserach: Theories and Issues,* Thousand Oaks: Sage.

Geurts, P. and Roosendaal, H. (2001). Estimating the direction of innovative change based on theory and mixed methods, *Quality and Quantity*, 35, 407-427.

Gorard, S. (1997a). *School Choice in an Established Market*, Aldershot: Ashgate.

Gorard, S. (1997b). A choice of methods: the methodology of choice, *Research in Education*, 57, 45-56.

Gorard, S. (2002a). Ethics and equity: pursuing the perspective of non-participants, *Social Research Update*, 39, 1-4.

Gorard, S. (2002b). Fostering scepticism: the importance of warranting claims, *Evaluation and Research in Education*, 16, 3, 136-149.

Gorard, S. (2002c). The role of causal models in education as a social science, *Evaluation and Research in Education*, 16, 1, 51-65.

Gorard, S. (2003). *Quantitative methods in social science: the role of numbers made easy*, London: Continuum.

Gorard, S. (2004a). Scepticism or clericalism? Theory as a barrier to combining methods, *Journal of Educational Enquiry*, 5, 1, 1-21.

Gorard, S. (2004b). Three abuses of 'theory': an engagement with Nash, *Journal of Educational Enquiry*, 5, 2, 19-29.

Gorard, S. (2006). Towards a judgement-based statistical analysis, *British Journal of Sociology of Education*, 27, 1, 67-80

Gorard, S. (2009). Measuring is more than assigning numbers, in Walford, G., Tucker, E, and Viswanathan, M. (Eds.). *Handbook of Measurement*, Sage, (forthcoming)..

Gorard, S. (2010). All evidence is equal: the flaw in statistical reasoning, *Oxford Review of Education*, (forthcoming)..

Gorard, S. and Rees, G. (2002). *Creating a learning society?*, Bristol: Policy Press.

Gorard, S., Rees, G. and Salisbury, J. (2001). The differential attainment of boys and girls at school: investigating the patterns and their determinants, *British Educational Research Journal*, 27, 2, 125-139.

Gorard, S., Roberts, K. and Taylor, C. (2004). What kind of creature is a design experiment?, *British Educational Research Journal*, 30, 4, 575-590.

Gorard, S., Taylor, C. & Fitz, J. (2003). *Schools, markets and choice policies*, London: Routledge.

Gorard, S., with Adnett, N., May, H., Slack, K., Smith, E. and Thomas, L. (2007). *Overcoming barriers to HE*, Stoke-on-Trent: Trentham Books.

Gorard, S., with Taylor, C. (2004). *Combining methods in educational and social research,* London: Open University Press.

Gray, J. and Densten, I. (1998). Integrating quantitative and qualitative analysis using latent and manifest variables, *Quality and Quantity*, 32, 419-431.

Guba, E. (1990). The alternative paradigm dialog, pp.17-27 in Guba, E. (Ed.). *The paradigm dialog*, London: Sage.

Hakim, C. (2000). *Research Design*, London: Routledge.

Hammersley, M. (2005). Countering the 'new orthodoxy' in educational research: a response to Phil Hodkinson, *British Educational Research Journal,* 31, 2, 139-155.

Johnson, R. and Onwuegbuzie, A. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come, *Educational Researcher*, 33, 7, 14-26.

Kuhn, T. (1970). *The structure of scientific revolutions*, Chicago: University of Chicago Press.

Lakatos, I. (1978). *The methodology of scientific research programmes*, Cambridge: Cambridge University Press.

MacKenzie, D. (1999). The zero-sum assumption, *Social Studies of Science*, 29, 2, 223-234.

Medical Research Council (2000). *A framework for development and evaluation of RCTs for complex interventions to improve health*, London: MRC.

Meehl, P. (1998). *The power of quantitative thinking*, speech delivered upon receipt of the James McKeen Cattell Fellow award at American Psychological Society, Washington DC, May 23rd.

Middleton, J., Gorard, S., Taylor, C. and Bannan-Ritland, B. (2008). *The 'compleat' design experiment: from soup to nuts*, pp. 21-46 in Kelly, A., Lesh, R., and Baek, J. (Eds.). *Handbook of Design Research Methods in Education: Innovations in Science, Technology, Engineering and Mathematic Learning and Teaching,* New York: Routledge

Nash, R. (2002). Numbers and narratives: further reflections in the sociology of education, *British Journal of Sociology of Education*, 23, 3, 397-412.

Onwuegbuzie, A. and Leech, N. (2005). Taking the "Q" out of research: teaching research methodology courses without the divide between quantitative and qualitative paradigms, Quality and Quantity, 38, 267-296.

Paul, J. and Marfo, K. (2001). Preparation of educational researchers in philosophical foundations of inquiry, *Review of Educational Research*, 71, 4, 525-547.

Perlesz, A. and Lindsay, J. (2003). Methodological triangulation in researching families: making sense of dissonant data, *International Journal of Social Research Methodology*, 6, 1, 25-40.

Quack theories (2002). *Russell Turpin's 'Characterization of quack theories'*, htttp://quasar.as.utexas.edu/billinfo/quack.html (accessed 9/5/02).

Robinson, D., Levin, J., Thomas, G., Pituch, K. and Vaughn, S. (2007). The incidence of 'causal' statements in teaching-and-learning research journals, *American Educational Research Journal,* 44, 2, 400-413.

Sale, J., Lohfeld, L. and Brazil, K. (2002). Revisiting the quantitative-qualitative debate: implications for mixed-methods research, *Quality and Quantity*, 36, 43-53.

Selwyn, N., Gorard, S. and Furlong, J. (2006). *Adult learning in the digital age*, London: RoutledgeFalmer.

Smith, E. and Gorard, S. (2005). 'They don't give us our marks': the role of formative feedback in student progress*, Assessment in Education*, 12, 1, 21-38.

Snow, C. (2001). Knowing what we know: children, teachers, researchers, *Educational Researcher*, 30, 7, 3-9.

Social Research Association (2003). Ethical guidelines, www.the-sra.org.uk (accessed 17/8/09).

Teddlie, C. and Tashakkori, A. (2003). Major issues and controversies in the usse do mixed methods, in Tashakkori, A. and Teddlie, C. (2003). *Handbook of mixed methods in social and behavioural research*, London: Sage.

White (2008). *Developing research questions: a guide for social scientists*, London: Palgrave.