SAGE researchmethods

Introducing Designs in the Cycle of Research

In: Research Design: Creating Robust Approaches for the Social Sciences

By: Stephen Gorard

Pub. Date: 2017
Access Date: August 29, 2018
Publishing Company: SAGE Publications, Inc.
City: 55 City Road
Print ISBN: 9781446249024
Online ISBN: 9781526431486
DOI: http://dx.doi.org/10.4135/9781526431486
Print pages: 13-21

Summary

- Different types of research have different purposes, from synthesising what we already know about a topic, through developing a new theory, to monitoring how well a new approach works in practice.
- Different research designs will be particularly suitable for these different types of research. No one design is always applicable or best suited for an entire programme of study.
- Designs can be classified in a number of ways, such as whether they involve a pre-specified comparison or not, whether there is a sequence of data collection episodes, and whether they involve a deliberate intervention or not.
- Research findings are intrinsically more convincing when the research design uses the 'better' or positive half of such classifications. For example, a comparative claim is intrinsically more convincing with a pre-specified set of comparator sub-groups.
- Such consideration reveals that, all other things being equal, something like a case study will always tend to be the least convincing design, while something like a randomised controlled trial will tend to be the most convincing.

Introduction

This chapter starts by looking at the concept of an over-arching research cycle, and suggests that different designs might be more or less appropriate at different phases in that cycle. Each design covered in this book, and many others, has a place in social science. This section tries to show how they fit together in a programme working towards the solution of a social science puzzle.

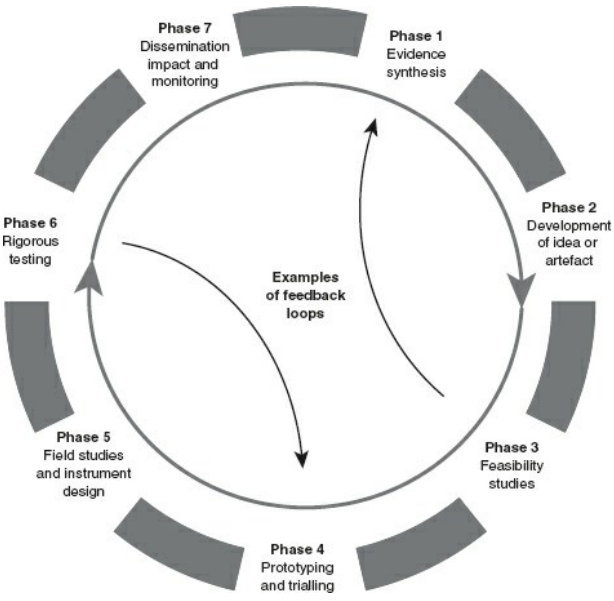
Research Cycle

Whatever area of social science you work in, it is likely that the field as a whole will look something like Figure 2.1. This is a simplified description of a full cycle for a research programme. 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 and Gorard 2007). The cycle is more properly a spiral which has no clear

beginning or end, in which phases overlap, can take place simultaneously, and could iterate almost endlessly.

The various phases illustrated should be recognisable to anyone working in areas of applied social science, like public policy. I discuss each of these steps in slightly more detail below, and then in succeeding chapters. Any one study or researcher might contribute to only some of these steps, but the field as a whole ought to progress in something like this fashion. Unfortunately, some fields seem to have got stuck in a limited range of phases (Gorard et al. 2011). Of course not all ideas come to fruition. There should be questions that are answered by a synthesis alone. Then some ideas should halt at a pilot stage where they show no promise. But equally, this should not always happen. Some ideas ought to be developed all of the way to become tested 'products' engineered into use for other researchers, policy-makers or practitioners. These products could be theories, practical protocols, policy interventions or genuine artefacts like software or training manuals. The key point for this book is that different designs will be more appropriate at different phases in the cycle. So a healthy field, as a whole, will use a wide range of research designs – which means of course that everyone needs to know something about all designs, even if only to be able to conduct an appropriately critical literature review in their field.

Figure 2.1 An Outline of the Full Cycle of Social Science Research and Development



Evidence Synthesis

The cycle might start with an evidence synthesis, which should use existing datasets and previously published literature in an unbiased way to produce a summary of what is already known in relation to the research question(s). The cycle might also end there, if existing

evidence answers the question satisfactorily. In practice it tends not to, partly because the external funding structure for research discourages such honesty (Chapter Thirteen), and partly because researchers are generally so poor at conducting a synthesis of existing evidence. The characteristics of a good synthesis are discussed in Chapter Three. Once the existing evidence synthesis is complete, it should be clear what remains unanswered in the area of interest, and this can lead to a definition of the problem to be solved, the research purpose, and research questions.

Development and Preparation

A high proportion of existing published work in social sciences seems to be in Phases 2 to 3 of the cycle. This is what may be envisaged as the research development phase. However, very little of this work currently starts with a serious attempt at research synthesis, offering instead only partial literature reviews and confusing conceptual frameworks. Even less of it moves on from this development work towards the preparation of a definitive large-scale study, or to creating something useful from the knowledge gained. Some of the reasons why this may be so were touched on in Chapter One.

Proto-typing and trialling of any new idea is vital, if resource is not to be wasted on a large-scale definitive study that has no chance of success. Trying ideas out at this development stage will tend to be done via small scale work, to minimise the risk and cost in case the idea does not work. With minimal risk and cost, several alternative ideas can be tried out in parallel. Feasibility studies can be as cheap as thought experiments (Chapter Nine), as simple as case studies, or they can be complex designs for multi-method data collection. Their purpose in the cycle is to assess the likelihood of success of the idea, and so to assist with the decision whether to proceed further in the cycle, or not. The decision can also be influenced by the potential gain from the idea, the resources required, and ethical issues.

Evaluation

Each prior phase might lead to a realisation that little more can be learnt and that the study is over, or that the programme needs radical revision and iteration to an earlier phase, or progression to a subsequent phase. However, the overall cycle can 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. 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. It is no good

knowing that an 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. Methods really are independent of design.

A simple Design Typology

The basic elements of research design introduced in Chapter One can be combined to form a wide variety of study designs. Most of these combinations would not have a well-known name but are as valid as any other combination nevertheless. The value of a design can only be judged in relation to the research questions it is intended to answer (Chapter Three). This in turn depends upon where the research programme is currently focused in terms of Figure 2.1. Some combinations of design elements have well-known names like ‘case study’ or ‘longitudinal’.

One way of classifying such standard designs is in terms of whether they are active or passive. An active design would include a controlled intervention, introduced as part of the study. Examples of active designs that you may already have heard of include randomised controlled trials (RCTs), and laboratory experiments. Quasi-experiments, action research, interrupted time series, regression discontinuity, and design studies may also involve a specific intervention. All of these are intrinsically more convincing in testing a causal claim than completely passive designs (Chapter Nine). A passive design may consider changes over time but these changes do not occur as part of the research itself, because there is no specific or controlled intervention. Examples include standard cohort research, other longitudinal designs, case studies, and comparative or cross-sectional approaches.

Another way of classifying designs could be in terms of whether they involve a pre-specified comparator group, or not (Chapter Seven). Obviously, comparative studies are intrinsically more convincing in testing a comparative claim than non-comparative ones. Examples include RCTs, natural experiments, and comparative research. Non-comparative designs include standard longitudinal approaches and case studies.

A third way of classifying designs could be in terms of whether they involve repeated measures or some other planned elapse of time between the start and a final measure, or not. Such

longitudinal designs are intrinsically more convincing in demonstrating a before-and-after claim than cross-sectional ones (Chapter Eight). Examples of designs with an automatic longitudinal element include standard cohort research, other longitudinal designs, RCTs, and natural experiments. Those usually without a longitudinal element include case studies, and comparative and cross-sectional approaches.

It is interesting that only RCTs and other experiments appear in the 'better' or positive half of each of these classifications. They are better for causal, comparative *and* time-dependent claims. In addition, RCTs have the advantage over quasi-experiments of having cases allocated to comparator groups at random. As shown in the rest of the book this means that given the right conditions and questions, an RCT or equivalent laboratory experiment is the best and most convincing design to use. Much of the rest of research, as in the full cycle, can be envisaged as working towards such a trial.

It is also notable that case studies are always in the worse half of each of these classifications. In themselves, they have no comparator, no intervention and no longitudinal element. As a design they are simply an episode of (possibly extended) data collection:

NO

This notation represents observations (O) or data collected at one point in time with no intervention, no pre-defined comparator groups, and so no rule about how cases were allocated to comparator groups. With so few of the elements of research design present here, there is little that such research can do beyond exploratory initial descriptive preparation for subsequent studies. Such work might be useful in generating ideas and possible explanations for a causal model perhaps. However, the authors of such work rarely seem to generate such ideas. Even where the authors of such work describe it as exploratory, they do not then progress with it to a later phase of the cycle. Instead they seem merely to move to another exploratory study (see Gorard et al. 2011 for examples).

A case study, in isolation, will never be the preferred design for any study that aims to be convincing or definitive. Case studies can be valuable, especially towards the start of the research cycle, largely because they are simple and quick to set up. I realise that some commentators would say that case studies are to be preferred because they allow a researcher to study a case in-depth. But you must recall the point from Chapter One. Designs are independent of the methods of data collection. It is as feasible that a case study was an examination of the financial accounts of one company as that a longitudinal study involved an in-depth observation of someone's adjustment to a new job over the first six months. No design has a monopoly on depth or breadth of data.

The need for warranted conclusions requires the researcher to identify the kind of claims to be made – such as descriptive, associative, or causal – and then ensure that the most appropriate possible design is used. A comparative claim *must* have an explicit and suitable comparator, although it is truly shocking how often this is not found. The warranting principle in Chapter Four 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 research-based claim should be the simplest explanation for the available evidence. What the research design should do is eliminate (or at least test or allow for) the greatest possible number of alternative explanations before the final claim to knowledge is made. In this way, the design eases the analysis process, and provides part of the warrant for the research claim. Design makes research better.

There are a host of already known designs other than those mentioned so far, and presumably many more waiting to be combined from the elements of research. Which design is used for any study should depend largely on the kind of claims and conclusions to be drawn. And these in turn depend on the research questions to be answered, which are the subject of the next chapter.

Exercises on Designs

1. Two doctoral researchers at a social science conference describe their research projects. The first researcher is a social historian, looking at the impact on the diet of agricultural labourers of the enclosure of common land in England. The researcher examines records and other sources relating to meals for the labourers in one area, from around 25 years before the recorded start of land enclosure, and then from around 50 years later. The second is an education researcher who has persuaded their own institution to schedule their adult education classes more flexibly, in order to encourage a wider range of participants. This researcher has conducted a survey of people taking adult education classes in the year before the change took place, and in the year after. The survey asked people about their occupational and educational backgrounds.
    a. What is the simplest design notation for each study?
    b. How would the design look different if the second researcher had used interviews with the adult learners, rather than a survey?
    c. Both researchers want to argue that there was a change from the first to the second episode of data collection, and that this change was due to enclosure in the first study and the rescheduling of classes in the second. What is the biggest problem for their argument?
    d. How would you advise them to re-design each study to try and overcome this problem?

2. A third researcher at the same conference has used quite a complex research design, involving multiple groups and episodes of data collection. A member of the audience asks what the design is called, and then mocks the presenter for not knowing. Is it reasonable for a researcher to describe a design but not know what it is called?
3. Why might an intervention that was found to be reasonably effective in Phase 6 of a research cycle turn out to be much less effective when rolled out into widespread practice? Come up with a range of suggestions, perhaps through discussion in a group.

Notes on Exercises on Designs

1. a. The simplest design notation for the first study could be:

```
N     O
N          [X]     O
```

Different lines are used here because the cases are different before and after the intervention of interest. Of course, the first researcher did not really intervene to create the mediaeval land enclosure system in England (how could they?). This is why I put the X in square brackets. But this limitation is the same for all history, retrospective accounts, archaeology, palaeontology and so on, and so I think it is reasonable to represent the design like this. The first researcher is looking at evidence from one time period and then looking at evidence from a later period, with an important intervening change. In some respects it does not matter whether this change has already happened or not, or quite how long ago it happened.

The simplest design notation for the second researcher could also be:

```
N     O
N          [X]     O
```

They have gathered evidence from one cohort of adult learners, waited until their institution altered something important, and then gathered evidence from a later cohort. In essence, the design is very similar to the first despite the surface differences in the topics.

b. If the second researcher used interviews rather than a survey, this would make no difference to the design. Data collection is only about what goes on within each O episode, and this is traditionally unspecified at the level of research design.

c. Both researchers face the same problem. They cannot tell whether any difference between the two time periods is due to the intervening specified change; whether it would have happened anyway; is due to differences between the people; or whether it

is just natural variation caused by incompleteness of records or the vagaries of data collection. Their argument would be unconvincing, as it stands.

d. Probably the simplest thing both researchers could do is to add a relevant comparison group to their design, to combine with the before-and-after element. The second researcher could look at the before-and-after participation rates in another institution that did not reschedule classes, or they could have rescheduled only some of their classes in the first year to see what happened. Therefore, the design could be:

```
N     O
N          [X]     O
N     O
N                 O
```

This design is discussed further as a difference-in-difference, in Chapter Ten. It is harder for the first researcher, but they could pick any of a variety of slightly weaker comparisons. For example, they could have looked at diets 50 years before the start of the study, or 50 years after it, and so tried to judge if the change specifically during the onset of land enclosure was remarkable in any way. Therefore, the design could be:

```
N     O
N          O
N               [X]     O
```

As you see, passive designs for causal questions can get complicated very quickly. This kind of design is discussed further, as an interrupted time series, in Chapter Ten.

2. It might be a little embarrassing for the third researcher if they were using what was clearly a simple and well-known design like a cohort study but had not learnt the name. Names are a useful shorthand as long as both the researcher and the audience mean the same thing by it. Not knowing the name would not affect the validity of the research of course. However, the implication here is that the design is not well-known. Perhaps it is some combination of two or more off-the-shelf designs. In this case it does not matter at all if it has no name. Even if the researcher gave it a name this would be useless for the presentation because it is then almost certain that the audience would not know what the name meant. Given that the design will have to be explained anyway the name might just be confusing. In general, knowing the names of designs is over-rated. Where the design is simple, such as a case study, there is often considerable confusion and disagreement between experts about what it means. It is therefore ambiguous and needs spelling out. Where the design is unfamiliar and complicated then the name will not lead to recognition and could be ambiguous. So, it will still need spelling out. Design notation, or similar, is more likely to clarify the design, and is safer than just using a name.
3. There are many reasons why an intervention that was found to be effective in Phase 6 of a

research cycle might turn out to be much less effective when rolled out into widespread practice. The results of the evaluation might have been faked or inadvertently misunderstood by the researchers involved (Chapter Four), perhaps because of a conflict of interest (Chapter Thirteen). Did you consider that? The intervention might have been better resourced as an experiment than when it was rolled out. This often happens. The evaluation might have involved only volunteers, whereas the rollout might be imposed on all, leading to sullen participation. The context for the evaluation might have been more suitable for success than the greater variety of settings encountered on rollout. The general training to use the intervention might be worse than it was during the experimental phase. And so on.

Suggested Reading

Gorard S. (2010). 'Research design, as independent of methods'. In **Teddlie C.**, and **Tashakkori A.** (eds) *Handbook of Mixed Methods.* Los Angeles: SAGE.

http://dx.doi.org/10.4135/9781526431486.n2