# Introduction

Imagine, for a moment, that you are a teacher and you hear about an exciting new approach that has been developed to enhance primary school children's reading skills. The approach has been organised into a clear programme that runs for one school term and includes a range of specific tasks and activities to be undertaken with the children on a whole-class basis. From what you can see, the programme looks convincing and certainly reflects the types of approaches you feel are important and that you wish to develop further with your own children. One of the first questions you will ask as a teacher will be: is this new programme likely to be effective? In particular, and before you risk adopting it with your own class, you will be keen to know if there is any evidence to show that it is likely to work better in terms of improving reading skills than the existing approaches you are using. You will also have in mind the particular characteristics and needs of your children when you are asking this question. You will thus also want to know: what are the chances that this programme would work for your particular class?

This book focuses on the research methods used in education to address precisely these types of questions. Starting with the basics, the book builds up your appreciation and knowledge of the methodological approaches used to test the effectiveness of a wide range of different educational programmes and interventions. It takes you through the process of how to design your study, given the particular type of programme or intervention you wish to evaluate, and then how to undertake that study. All of the key issues involved in the collection and analysis of the data and the writing

up of findings are covered in detail. In working through this book you will gain the knowledge and skills required to design and undertake a piece of research that will provide you with strong and robust evidence of whether a particular educational programme or intervention works or not. Even if you do not intend to conduct such research yourself, this book will provide you with invaluable insights into, and an appreciation of, the appropriate methods to use to evaluate the effectiveness of educational programmes and thus the skills you will need to critically assess the strength of the evidence that you will read elsewhere.

## Working out what works

Of course, each child is different. What might be effective for one child in your class may not be effective for another. Also, each class and school is different. An approach to reading that may work well, overall, with your children may not work well, or may even have a negative effect, with another class in another school. These questions about what works are therefore, inevitably, questions about general tendencies. It is simply not possible, through research, to predict with certainty that a particular approach to reading will definitely lead to additional gains in reading skills for a specific child or a particular class of children. Rather, all we can do is seek evidence of what is likely to work, for what types of children and in what contexts. Turning back to this exciting new reading programme you have heard about, you may find that there have been several rigorous studies published to date that have all shown, to varying degrees, that this new programme tends to increase children's reading scores by the equivalent of about an additional three months' progress compared to those of children who just carried on as normal. These studies should give you sufficient encouragement and confidence to try the programme out with your own class. However, you will never know whether it will work with your children until you actually try it. Moreover, even if you find it does work, it is unlikely to have the same effect for all of your children. While your class may improve their reading scores by an additional three months' progress compared to what they would have done otherwise, this will only be the average level of improvement. By definition, and being the average, we know that half the children will have made greater gains than this but that the other half will have made lesser gains. Indeed some may have even regressed over this period.

The key point therefore is that research can only ever tell us how particular groups of children are likely to respond, *on average*, to a specific educational programme or intervention. However, the availability of a

body of rigorous evidence that can provide teachers with insights of this kind for a wide range of differing educational programmes and interventions is undoubtedly extremely valuable. It is precisely this question of how we can generate robust evidence of whether an educational programme or intervention is effective for a particular group of children that provides the focus for this book. Methodologically, there are a number of clear steps we will need to take if we want to design a research study capable of addressing this particular question in a rigorous and convincing way. As an illustration, let us return to the example of the new reading programme for primary children that has been designed to be delivered on a whole-class basis over one school term. First of all we need to gather some objective data on how much progress the children made with their reading over the duration of the term when they were engaged with the new programme. To do this we could use any one of a number of standardised reading tests at the start of the term and then again at the end to get some measure of how much progress has been made by the children on average. However, we have to take into consideration that they are likely to make some reading progress during this period regardless of the teaching method used, simply due to attending school. As such, if we want to estimate what the *additional* effects are, on average, of introducing this new reading programme, we also need to see how a group of similar children progress over the same period who were not using that approach in class. This is our second step – introducing a comparison, or control group. Ultimately, if the children using the new approach make greater gains in their reading skills compared to those in the comparison group then this would provide us with evidence that the new approach is likely to be more effective, on average, for those particular children. Within this, we can also extend our analysis of the data to compare the progress made by boys and girls separately in each school, and by any other factor that we feel might have an influence on the effectiveness of the programme on average.

The main reason we are able to draw robust conclusions regarding the effectiveness of this reading programme from this type of study design is that we are assuming that the two groups of children are similar. Thus, and to put it simply, if the children in the schools using the new reading scheme are broadly the same as the children in the schools carrying on as normal then the only systematic difference between them is the fact that one group is using the new reading programme while the other is not. If this is the case then any differences in the reading progress made by the two groups of children must be due to the effects of the reading programme, as the two groups are the same in all other respects and are thus subject to the

same range of other external social processes and factors within and beyond the school. Of course, this assumption that the two groups are the same is a critical one. It is for this reason that the third step in the research design process is that we need to create two matched groups, and we need to do this through random allocation. More specifically, and in the case of this new reading programme, we could work with maybe 40 or more primary schools and randomly select half of them to run the new programme and the other half to continue as normal. The crucial point to note is that by randomly splitting the schools into two groups we have maximised the likelihood that we have two matched groups. Through the randomisation process, and with a sufficiently large sample, we can be fairly confident in assuming that all of the other myriad of factors that will impact upon children's progress in reading will be evenly balanced across the two groups.

What has been described above is, in essence, a randomised controlled trial or RCT. It is a *trial* of a particular educational programme or intervention to assess whether it is effective; it is a *controlled* trial because it compares the progress made by those children taking the programme or intervention with a comparison or control group of children who do not and who continue as normal; and it is *randomised* because the children have been randomly allocated to the two groups being compared. It is this design that provides the focus for this present book. In the chapters to follow we will look at the broader philosophical and methodological debates regarding the use of RCTs in social and educational research as well as setting out, in some detail, the practical issues involved in choosing the right design for the particular programme or intervention you wish to evaluate. The book also takes you through, step-by-step, how to collect and analyse the data from an RCT and how to report it thoroughly.

## RCTs: What works and ongoing controversies

Over the last ten years we, the authors, have collectively run over 30 RCTs in education. To date, these have involved over 800 schools and other educational settings across Northern Ireland, the Republic of Ireland, England and Scotland. Each one has involved working collaboratively and in partnership with a wide range of schools and organisations involved in developing and delivering educational programmes and interventions. Our experience over this time is that the type of approach described above is largely uncontroversial. Put simply, teachers, pupils, parents and other key stakeholders just seem to 'get it'. They understand the need for a control group and, with a little explanation, they soon come to fully appreciate the importance of randomisation. Perhaps most importantly,

they are all passionately concerned with the effectiveness of what they do and whether, ultimately, it is having a clear and measurable impact on children's learning and development. This is not to say that the process of undertaking an RCT is easy, or that it is without complications. As will be described throughout this book, running RCTs in education presents many challenges and dilemmas. However, these challenges – of research design, of ethics, of negotiating and maintaining relationships with stakeholders, of analysing, interpreting and presenting findings – are ones faced by all social and educational researchers, regardless of the particular methods they employ. The point about RCTs being uncontroversial is that we have simply not encountered any resistance to the use of trials in education or faced any unique problems with regard to running trials compared to any other type of research.

We make this point about the uncontroversial nature of RCTs *in practice* because they have faced, and continue to face, considerable resistance to their use in education *in theory*, particularly from significant sections of the education research community. Unfortunately, such opposition to RCTs appears to be embedded in the culture of the academic community in education. Take, for example, one of the leading textbooks *Research Methods in Education*, written by Cohen, Manion and Morrison and now in its seventh edition. It is arguably the most popular and widely used methodology textbook in education in the UK. This is how they present RCTs to students:

> This mode of an experiment, premised on notions of randomization, isolation and control of variables in order to establish causality, may be appropriate for a laboratory, though whether, in fact, a social situation either ever *could become* the antiseptic, artificial world of a laboratory or *should become* such a world is both an empirical and a moral question respectively … Further, the ethical dilemmas of treating humans as manipulable, controllable and inanimate are considerable … [T]he experimental approach may be fundamentally flawed in assuming that a single cause produces an effect. Further, it may be that the setting effects are acting causally, rather than the intervention itself.
>
> […]
>
> Schools and classrooms are not the antiseptic, reductionist, analysed-out or analysable-out world of the laboratory … Generalizability from the laboratory to the classroom is dangerous, yet with field experiments, with their loss of control of variables, generalizability might be equally dangerous … Randomized controlled trials belong to a discredited view of science as positivism.
>
> (Cohen et al., 2011: 314 and 318, original emphases)

The tone of this criticism is notable, as is the complete mismatch between these trenchant criticisms of RCTs *in theory* and the reality of their use in a wide range of educational settings *in practice*. Indeed, the authors make clear to students reading their textbook – who will become the next generation of education researchers – that 'RCTs are often not possible in education' (Cohen et al., 2011: 66), even though there have been many thousands of RCTs conducted in education successfully to date (Oakley, 2006). Given this mismatch between theory and practice, it is difficult not to agree with Oakley (2006) that much of the critique betrays a fundamental, worrying and widespread ignorance among significant sections of the educational research community of the use of RCTs in reality. It is certainly our experience that the 'constant recycling of stylised objections' to the use of RCTs and the associated failure to recognise or engage with the many detailed discussions that have taken place in the methodological literature regarding these, as Oakley (2006: 69) noted a decade ago, is still commonplace among the educational research community today. Indeed, the polarised nature of the debate over these last ten years, and the relative lack of any real and constructive dialogue between advocates of and objectors to RCTs, would suggest that we have witnessed the resumption of the classic paradigm wars, this time being waged on the site created by the emergence of evidence-based practice (Hammersley, 2008).

## About this book

This book will debunk many of the myths and arguments surrounding RCTs as touched upon above. By taking you carefully through the realities of conducting RCTs in practice in education, the book will demonstrate very clearly that the types of criticisms listed above are not just mistaken but are also ill-founded and that they do, unfortunately, betray a fundamental lack of understanding of the use of RCTs in practice. It will be seen, for example, that far from RCTs being done *on* teachers, there are many models of collaborative practice where teachers, pupils and other key stakeholders are centrally involved in the design of trials and in the analysis, interpretation and presentation of findings. Also, the continually recycled concerns about ethics tend to betray a lack of awareness of the many RCTs where no child is denied the programme in question; it is just that there is a delay before the control group are delivered it. In addition, rather than RCTs ignoring contexts and being unable to undertake within-programme analyses, it is quite possible and indeed increasingly commonplace for trials to pay close attention in their design and analysis

to the particular contexts within which a programme is delivered, how its delivery varies between contexts, and how – within any one context – the effectiveness of the programme may vary systematically between differing sub-groups of learners. Moreover, many RCTs are multi-method in design and include a parallel qualitative component precisely to help understand the quantitative findings that emerge from the trial, and thus to address directly the questions of *why* a particular programme has (or has not) been effective and/or why its effectiveness has varied between contexts or groups of learners.

Given the widespread misunderstanding of and resistance to RCTs in education, we will therefore begin, in Chapter 1, by looking in more detail at some of the core methodological, philosophical, ethical and political debates touched upon above. The chapter sets out the rationale for RCTs and also the key criticisms levelled at their use in education and social research. These criticisms are each addressed, in turn, with reference to real-life examples of RCTs that we have conducted in practice. Through this, the chapter then sets out a more considered philosophical position for RCTs in education and social research, drawing upon critical realism, that provides a framework for the rest of the book.

Chapters 2 and 3 then move on to look at issues relating to the design of RCTs. Chapter 2 encourages you to take an initial step back from the educational programme or intervention you may be interested in so that you can map out and understand the theory of change that underpins it. The chapter will set out a practical framework for doing this through the use of a technique known as logic modelling. It will be shown how logic modelling is an invaluable tool for helping to think through the theory of change and to identify what the anticipated outcomes are for a particular programme or intervention and precisely how these are expected to be achieved. The chapter then looks, in some detail, at how to choose appropriate measures for the outcomes specified. With the wider theory of change in mind, as set out by the logic model and associated anticipated outcomes, the book then moves on, in Chapter 3, to consider the key issues involved in the design of an RCT. This chapter sets out clearly a range of common designs and which ones are most appropriate to use given the nature of the educational programme or intervention under consideration. Within this, the chapter provides step-by-step advice on how to calculate the size of the sample needed for the RCT, and also how to incorporate a qualitative component to study the process of delivery of the programme or intervention, as well as to help understand the findings that arise from the quantitative trial data.

The next two chapters focus specifically on how to analyse the data gathered from an RCT. While the importance of incorporating a strong qualitative component to the design of an RCT has been stressed, because of the limits of space, these two chapters will focus specifically on the appropriate methods for analysing the quantitative trial data. Chapter 4 begins this by setting out the overall approach to analysing a simple RCT (i.e. when individual children have been randomly allocated to the control and intervention groups). The chapter will take you through the steps involved in determining the overall effects of the programme or intervention concerned and also how to ascertain whether it has been more or less effective for particular sub-groups within the sample. Chapter 5 then extends this general framework by setting out how to undertake these analyses when you have different types of outcome data and also when the data are gathered from a cluster randomised controlled trial (i.e. when the study has involved randomly allocating whole classes or schools to either the control or intervention groups).

The final chapter, Chapter 6, focuses on how to write up and report the findings of randomised trials. Taking the CONSORT statement as the guide, the chapter will set out clearly the information that needs to be reported and in what format. It will also provide you with an appreciation of how the findings of individual RCTs are used in the synthesis of research evidence through the use of systematic reviews and meta-analysis. In all of this, the chapter will stress the need to be tentative in interpreting the findings as well as the claims that are made from these.

## Prior knowledge and suggested background reading

This book does not assume any prior knowledge of evaluation methods in general or of RCTs in particular. It has been specifically written for the generalist educational or social researcher who wishes to extend their knowledge and skills to the use of RCTs. In this sense, and as outlined above, the book is designed to take you through each stage of the process of designing, undertaking and writing up an RCT in a clear step-by-step manner. However, it does assume some basic understanding and appreciation of research methods in two respects. Firstly, it is assumed that you already have a good knowledge of qualitative methods and also case study methods. Their importance in the context of an RCT will be stressed in the book and how they can be used will be set out in

Chapter 3. However, and because of the limits of space, the book will not cover the key issues involved in undertaking case studies or in using particular qualitative methods such as semi-structured interviewing, focus group discussions and participant observation. Secondly, it is assumed that you are also familiar with the statistical software package SPSS and that you have some prior knowledge and understanding of basic statistics. In particular, it is assumed that you will be familiar with the key concepts associated with descriptive statistics – namely percentages, means and standard deviations and their graphical display in bar charts and boxplots. It is also assumed that you have some appreciation of inferential statistics and, in particular, of the notion of statistical significance and the associated use of 'p values' and confidence intervals.

If you do not have a sufficient grasp of qualitative methods or basic statistics then you will find it beneficial to consult one or more of the references set out below before reading this book.

## Further Reading

### Qualitative Methods

*Qualitative Research in Education*, by Liz Atkins and Susan Wallace (Sage Publications, 2012).

*Research Methods for Education*, 2nd edn, by Peter Newby (Routledge, 2014).

*Introduction to Research Methods in Education*, 2nd edn, by Keith Punch and Alis Oancea (Sage Publications, 2015).

### Case Studies

*Using Case Study in Education Research*, by Lorna Hamilton and Connie Corbett-Whittier (Sage Publications, 2013).

*The Anatomy of the Case Study*, by Gary Thomas and Kevin Myers (Sage Publications, 2015).

*Case Study Research: Design and Methods*, 5th edn, by Robert Yin (Sage Publications, 2013).

### Statistics

*Statistics: A Gentle Introduction*, 3rd edn, by Frederick Coolidge (Sage Publications, 2013).

*Beginning Statistics: An Introduction for Social Scientists*, 2nd edn, by Liam Foster, Ian Diamond and Julie Jefferies (Sage Publications, 2014).

*Statistics Explained*, 3rd edn, by Perry Hinton (Routledge, 2014).

## Using SPSS

*SPSS Explained*, 2nd edn, by Perry Hinton, Isabella McMurray and Charlotte Brownlow (Routledge, 2014).

*SPSS Survival Manual*, 5th edn, by Julie Pallant (McGraw Hill Education, 2013).

*Discovering Statistics Using IBM SPSS Statistics*, 4th edn, by Andy Field (Sage Publications, 2013).