Introduction

The state of educational research is widely agreed to be disappointing. It is variously criticised for indulging itself in airy theory rather than the practicalities of the classroom, for being slow to emulate the theoretical advances of other disciplinary fields, for not focusing on 'what works' in schools, for leaving theory and practice in isolation from each other, for failing to match the achievements of medical research, for being insufficiently rigorous, and no doubt for numerous other sins. It will immediately be obvious that these criticisms do not all pull in the same direction.

While we too are disappointed by the current and recent condition of educational research, this is for rather different reasons than most of the above. We think that many of the problems of educational research stem from its frequent attempts to imitate scientific research, and especially the randomised controlled trials (RCTs) that are often considered the 'gold standard' in medical research; from an ill-defined obsession with criteria such as 'rigour' and 'robustness'; and, above all – and like so much here no doubt following on from the emulation of science – from the general sense that proper research is primarily or exclusively empirical. The good researcher, on this model of research, spends her time (pausing only to take a 'research methods course') out and about gathering data, rather than in reading, theoretical analysis and reflection.

Here are some symptoms of these assumptions and tendencies. Both are real examples, but for obvious reasons we do not include details that could identify anyone or any particular institution. First, a candidate being interviewed for a Professorship in a social science discipline (not Education) in an eminent UK university. At one point the Chair of the Appointing Committee asked: 'You have told us all about your ideas and theories. Now what about actual research?' The implication here seems to be that real research consists in something other than developing ideas, thinking and writing: it must be more a matter of going out and finding. Research is empirical, or it is not really (not 'actual') research.

Second, an undergraduate student whose proposed topic for his dissertation was self-esteem and the Boy Scout movement (a composite example). He hoped to establish that Scouting improves self-esteem (he is himself an enthusiastic ex-Scout and now Scout Leader). His proposal consisted mainly of interview questions, and it was on these that he sought advice. He intended to ask a representative range of Scouts whether they believed their membership of the organisation had improved their self-esteem. On a scale of one to ten, where did they think they were, in terms of self-esteem, before they joined, and where did they think they were now? What particular activities had most improved their self-esteem? And so on. The point here is not to criticise this student. It was not difficult to persuade him that there were difficulties here, causation being one (it is hard to show that what someone thinks has increased their self-esteem is in fact the cause of that increase) but by far the more complex difficulty being the idea of self-esteem (see Chapter 11). He had taken a compulsory module on 'research methods' which had persuaded him that although research might include conceptual problems, generally to be solved by clarity of definition, the real business lay in interviews, questionnaires, coding and analysis of both, and establishing the validity of conclusions, which would look roughly like this: '67 per cent of my respondents agreed that Scouting had improved their self-esteem either to some extent or to a great extent. However, completed questionnaires were returned by only nine people, i.e., less than 20 per cent of those to whom they were sent, and it proved possible to interview fewer Scouts than originally hoped. More research is needed into this important question' (the wording here is adapted from several relevantly similar dissertations).

It is important to make the point, crude though in many ways it is, that educational research, like many other forms of academic research, is demonstrably done in the context of particular material circumstances that at least in part determine the topics we address, the styles in which we write, the disciplines from which we write and so on. There is, above all, the formal, regular evaluation of academic research, which in the UK has taken place over some 20 years under the aegis of several Research Assessment Exercises (RAEs) and now operates under what is designated as the Research Excellence Framework (REF). If we do not say as much about this in detail as could be said that is partly to avoid being parochial because the context is specific to the UK, although similar exercises are now to be found in many other countries, and partly because (one hopes) the whole thing will before long come to seem a minor curiosity, one more sign of the absurd performative culture of our own time. However, it is worth touching on a number of aspects of the REF that exemplify the

tendency towards empiricism and scientism – that is, excessive respect for the image and tropes of science – that we are here criticising.

First, in this Exercise or Framework the subject area of education is treated as if its model were psychology. A relatively trivial effect of this is that academic journals in the social sciences now widely require submissions to follow the referencing and other conventions of the American Psychological Association (APA). A more important effect is that journal articles are taken more seriously as contenders for high ranking than book chapters or even whole books. The reason is that in psychology there is a general assumption that cutting-edge research will appear in the latest issue of a prominent journal, while a book will essentially be either an exercise in popularisation or a textbook that summarises the current research orthodoxy rather than disseminating the latest research findings. A new discovery could not wait (so the assumption runs) for the lengthy process of bringing a book to publication. (Medical research hovers in the background here as the exemplar.) However, it is of course not at all clear that educational research is likely to come up with dramatic 'findings' of a comparable kind. A thoughtful overview, at book length, of the idea of the university or the place of religious education in schools, for instance, might be both important and necessary. Another effect is that new researchers in education are encouraged to join established 'research teams' and become one of the many authors of a journal article, as in psychology: as if educational research was cumulative and piecemeal. There is a case for saying that the educational researcher, by contrast, needs to develop a voice and a perspective of her own. It is important to add that psychology itself as an academic subject is steadily being driven towards the more 'scientific' end of the subject by the febrile, and not necessarily realistic, expectation that a richer understanding of human beings is promised by analysis of DNA, brain scans and neurobiology. Specialists in these fields are now beginning to be employed in university departments of education themselves.

Second, the impact requirement that has made its first appearance in the REF evaluation to be undertaken in 2014 means that an academic department must include in its submission some examples (roughly one for every ten active researchers) of how its research benefits the wider, and specifically non-academic, community. Much could be said about this requirement and about the assumptions being made, for instance, about how one is to predict such benefits; or about how a department of Philosophy or English Literature or Pure Mathematics is supposed to be able to show that it has – strictly speaking, that it is going to have – 'impact'. Some of the absurdities of 'impact', and the way it is made to

seem reasonable by some of the tropes of a debased notion of science, are well brought out by Fred Inglis (2012) in a *Times Higher Education* polemic:

> It is believed by the helots of the Department for Business, Industry and Skills that, like everything else in the world, impact is measurable by number. So if you are four-star REFable in, say, medical studies, then to appear in *The Lancet* is the measure of fame, and the journal is so canonised because everybody cites it; contributors, moreover, cite themselves, and the consequent citation index gives *The Lancet* a score of 33.63, whereas *Medical Teacher*, an honoured journal much favoured by GPs, struggles along, starved of citation, at 1.494. It is relevant to add that *The Lancet* score rocketed upwards in terms of its citations when, in a picturesque example, it published in good faith research claiming to discover a causal link between autism and MMR injections in babyhood. Numberless citations in repudiation of certainly unmethodical and probably dishonest work contributed largely to the journal's lordly score.

Even in science itself the criterion of impact is dubious: when applied, as Inglis notes, to the place of political ideas in historical explanation, or to 'a serious and careful book on the importance of beauty as a concept central to a decent education for children' which 'is rejected by some zombie evaluator as not recognisable as research at all' (ibid), the nonsense of treating all research as if it was like discovering a brilliant new technique for making artificial hip joints should be plain.

The confluence of a number of circumstances has led to what we might think of as the autonomy of quasi-scientific technique in research. Students of Physics, Biology or Chemistry need to learn laboratory techniques (using an oscilloscope; microscopy, staining and culture techniques; recrystallisation and melting point determination), so it is a small step to the idea that all knowledge and understanding require the same approach. This dovetails neatly with the longstanding assumption that there are such things as research methods (an odder assumption than it may appear, see Chapter 3) that can readily be acquired. The general tendency here has been further fed, as far as educational research is concerned, by the attraction of recruiting overseas fee-paying students whose command of English, although impressive, does not enable them readily to engage with conceptual or philosophical dimensions of education. Different techniques for collecting data for empirical 'research projects', however, are relatively easy to grasp, not least because the quicker you can get to numbers and quantities the quicker you leave behind the tricky problems of concepts and philosophical ideas. T-test, ANOVA, ANCOVA, chi-square, linear regression, factor analysis and a host of other techniques have the further virtue of making educational research look 'scientific'. We call these 'autonomous' techniques because

they tend to be presented as if they can be learned independently of any particular area of enquiry and then bolted on to a wide range of social science projects, whether they are nominally concerned with education, use of new technologies or substance abuse. Clearly research methodology is the specialism to have these days, as an internet investigation of the research interests claimed by academic educationists will readily confirm.

Yet another sign of how educational research is driven by contingent circumstances that favour empiricism is what is sometimes called 'sponsorism'. This is the designing of research in the light of the agenda of funding bodies. Where once a researcher might have found an interesting field of enquiry and then turned her attention to possible sources of funding, matters are now often the other way round. Funding, of course, is typically used for the collection and analysis of data, and for the appointment of researchers for that purpose: it is seldom awarded to give an academic the space and time in which to think. Thus the imperative is inexorably towards certain kinds of educational research and away from others. As a result – both the consequence of sponsorism and a further incentive for it – it is now a prerequisite for the passing of probation or for promotion at many universities that one should at least have made applications for research funding (naturally this means that funding bodies are deluged with applications so that actual funding is more and more difficult to win). New academic colleagues, we find, simply take for granted that the search for funding is their first research task and react to the older idea – that an interesting question might take priority – with discomfort and surprise. They tend to observe that 'this is the game' that they have to play: a response that reveals just how much educational research is now being driven by considerations other than the generation of knowledge and understanding.

A particular example may be helpful at this point. The European Educational Research Quality Indicators (EERQI) project promises a thoroughly scientific approach to evaluating research on education. It was 'motivated by the fact that the international notion of scientific quality as being the main determinant on which research is funded and supported' is unreliable. It will 'develop new indicators and methodologies to determine quality of educational research publications' (from the EERQI website). The 'prototype framework' involves 'bibliometric analysis and reference linking, semantic and linguistic analysis of full texts and citations, text strings and metadata correlations' (from Flyer No. 1, Promoting Research Quality). The 'new technological possibilities provided by natural language processing tools for content analysis

and for text mining of digitally-available scientific documents' (EERQI website) offer to replace human judgement with scientific, computerised analysis.

No doubt the general allure of 'science' has much to answer for in all this. One notes the talk of 'scientific quality', 'methodologies', 'metadata', 'technological possibilities' and so on, all adding to the rhetoric that suggests something very precise and impressive is being generated here. The quality of a research article in education, EERQI promises, will before long be established electronically (what a blessing this would be for the university Research Manager looking for 'objective' performance indicators). No doubt the software will favour work that is written 'scientifically', that is employing 'text strings' that draw on the terminology of science and those parts of social science, such as randomised controlled trials, that echo it. Researchers will no doubt learn to play the system, incorporating in their writing the key phrases and tropes. Thus, as with sponsorism, the anticipated reception of research will be prioritised over all other considerations.

EERQI offers the possibility of replacing human judgement and interpretation with what is in essence a sophisticated algorithm. There are wider social and cultural tendencies in play here. Oancea and Pring (2008: 25) plausibly suggest that politicians, desperate for 'certainty and solid foundations for their policies', confuse scientific evidence with proof. Proof is presumably seen as unconditionally compelling and as absolving the policy-maker of responsibility if the policy comes unstuck. It may seem to have the additional and very welcome attraction of releasing the policy-maker from having to use his or her judgement. Judgement, as Oancea and Pring also imply, is worrying to an age that tends to suppose that there is little space between scientific proof on the one hand and mere subjectivity on the other. We return to the discussion of judgement elsewhere in this book (see Chapters 3, 8 and 9).

We might, then, criticise much of the more crudely empirical and supposedly 'scientific' educational research on the basis of some of its contingent limitations, such as the readiness to collude with national (and, increasingly, international) protocols for measuring research quality (such as the Research Excellence Framework in the UK or the What Works Clearinghouse in the US), and the increasing tendency to prioritise funding possibilities over the intrinsic interest or intellectual richness of the problem or topic being researched. These contingent limitations are interesting, we believe, because they unsettle any notion that the empirical or 'scientific' turn in educational research is somehow inevitable or a move in the direction of progress. We have more to say about

the genealogy of these notions – that is, the particularity of the historical circumstances from which they have arisen – later in the book (see especially Chapter 3).

For the most part, however, we follow a different direction in this book. This direction can be indicated by sketching the important distinction between *erklärung* and *verstehen*. We shall return to it many times in these pages. It is the distinction between explanation of the kind familiar to science (*erklärung*) and understanding (*verstehen*) of a less technical kind. A simple example of *erklärung* might go thus: you discover a small puddle of water in your living room. You set about trying to discover the source: it might be a leak in the piping among the floorboards of the room above, water entering from a leaking drainpipe, the result of someone watering a pot-plant carelessly or the cat having an accident. If it is caused by one of these things it will almost certainly not have been caused by one of the others. Your investigation proceeds in the manner of shedding light (the root of the word *erklärung*), literally so perhaps if you take up the floorboards. Empirical investigation is the right and proper approach and will be most effective if you have sufficient light to see what has happened.

By contrast, suppose that you have a colleague who sits through meetings with a frozen expression, whose teaching students begin to complain is lacklustre, whom you several times find with her head in her hands at her desk. You might reasonably wonder if she is suffering from depression, or losing her professional motivation, or experiencing the symptoms of overwork. It might of course be a combination of all of these factors, or more. There is no clear distinction between depression and demotivation. You do not bring investigative techniques to bear, with any expectation that you will at some point find the 'true answer'. This is rather the domain of *verstehen*. You employ your ordinary human understanding to try to make sense of your colleague's state – for all that there is an increasing tendency to suppose brain-imaging or some new scientific technique will come up with the explanation. Here what you are attempting to do is find the *meaning* of your colleague's behaviour. For this you need to have some understanding of what it is like to feel depressed. Thus it is helpful that most of us have been unhappy at various points in our lives, if not 'clinically depressed', yet at the same time there is the danger of imagining that our colleague's life is following the pattern of our own ('Ah, I can remember how low I became when I was turned down for promotion – I know just how you feel!'). Things are made more difficult by the fact that if your colleague agrees with your 'diagnosis' it does not necessarily mean that you were right: people can be persuaded to accede to the views of others whom they esteem (psychoanalysts are particularly aware of this

problem). Nor does it mean that you were wrong if your colleague vehemently rejects what you say: you might reasonably wonder if this shows just how close to the painful truth you have come.

We shall be trying to show in all sorts of ways in this book that 'understanding education and educational research' requires – can hardly proceed without – *verstehen*. An extended example may help to clarify the argument at this point. Consider a lecturer who is teaching classes to undergraduates on the idea of education: who is helping them to understand education, to make sense of it. She wants them to see that there may be more to education than just passing exams and tests and achieving good marks on essays: that there is a tradition (or more accurately a number of overlapping traditions) in which education is a matter of encouraging people to think for themselves, of expanding their horizons, of helping them to see through and become liberated from the common assumptions of their time. All of this, for a lecturer in England, has an added dimension at the moment since English students are now incurring significant debts by going to university; the government is prompting them to think of themselves as consumers, in the hope that they will make demands of their teachers that will 'drive up' standards of university teaching. The lecturer wants her students to see through some of the absurdity of this – especially the whole neo-liberal language of 'the market' in which it is put. She wants them to make demands, but she wants those demands to be the right demands: that they be challenged, perhaps, as long as they are supported in facing these intellectual challenges. While the lecturer needs to see if they have understood what she has taught them, she can hardly rely on a simple test of any sort for this, partly because the ideas involved are too complex and partly because there would be a horrible paradox in trying to liberate them from the idea that education is all about tests by subjecting them to a test. For this and other reasons, she cannot therefore determine whether they have 'achieved the outcomes' of the course in the way a medical researcher might establish that a particular drug has cured their headache. That is partly because in an odd way she wants them to go on suffering from the headache: to feel the pull and the temptation of 'the market' in education while still struggling to free themselves from it. If a student told her at the end of the second lecture that she now understood perfectly clearly that the good of education could not be thought of as a consumer good, and listed four good reasons for saying so, she would be uneasy that the student was merely repeating ideas from her lecture as if they were orthodoxy, without any real engagement with them. It will certainly be no good to discover, on the basis of a short questionnaire, whether the

students are 'satisfied' with her teaching, because she wants them to be both satisfied ('Yes, this is interesting!') and dissatisfied ('I think there is something wrong in your argument here. In your lecture you said …').

How, then, does she discover whether her teaching is any good – whether she is indeed helping her students to make sense of education? She has to interpret the students' responses – to understand, in the spirit of *verstehen*, what the responses mean or amount to. The last comment above ('I think there is something wrong in your argument here. In your lecture you said …') could be made in the offended tone of one who has spotted an elementary contradiction that looks like the sign of teaching that has not been properly prepared, or in the excited tone of one who is beginning to sense her capacity for thinking for herself. The student who complains that the recommended reading is too hard may prompt the lecturer to reflect that she is indeed expecting first-years to engage with material better set for second- or third-years; on the other hand she may respond by gently suggesting that, like other complex academic articles, this particular one will not yield its meaning to a single (and perhaps cursory) reading – or that, again in the spirit of *verstehen*, it is not to be thought of as having just one meaning in the way that there is normally just one explanation for the puddle on the floor. She interprets the student's response in the light of how other students are responding and have responded in the past. She interprets the words of this student, here and now: someone who through encounters in seminars has already struck her as someone to take seriously, or as someone who needs to be told to go away and read the article properly. She may well revise her judgement of the student later, just as she may reconsider her own response, deciding that she has been too accommodating, impatient or brusque.

It makes perfectly good sense to say that this lecturer and her students are *reading* each other. As we say in ordinary English: 'how did you read his behaviour at that meeting?' – meaning, what did you make of it, how did you interpret it? The reading here is not quick: patience is necessary, and so is a continuous scrutinising on the lecturer's part of the quality of her 'reading' so far. She has to be alert to the features of *these* students, rather than seeing them as just students, indifferently: as if she were to imagine that all first-year students are reluctant to do the reading set for them. Here too is an analogy with reading: the good reader attends to the specific poem in front of her, rather than importing fantasies about poems in general. She attunes herself to it rather than bringing to bear a rigid (and, notice, 'rigorous') model or method for interpreting all poems indifferently. With both text and students it is important that we bring

ourselves to it or them: that we are ready to be personally disconcerted, caught off-guard. This might be by a moving line in the poem. It might be by something said in a seminar. One of us recalls, in a discussion with students of why they had come to university and what they thought university was for, a student who said, quietly, 'that is what I came to university to find out'. Here was a comment and a moment that needed to be given space, to be met with the same quiet thoughtfulness with which it was uttered.

It is the analogy between reading text and reading people that leads us to think of interpretation and the *verstehen* dimension of social science as involving the kind of qualities that Simone Weil (1951) writes of when she tells us that the ability to attend, to give people or texts (or God, in her distinctive way of thinking, but it is not necessary to share Weil's religious views here) one's attention, is rarer and more valuable than we generally realise. Patience, quietness, openness, sensitivity: these are some of its dimensions. They take us a long way from the determinedly 'scientific' and 'rigorous' (i.e., inflexible) outlook that characterises so much educational research these days, and a long way from the frenzied rushing around that substitutes for intelligent and responsive thought. *Waiting*: as we might say, when asked what we make of this poem or that person, 'I need to read it again before I can answer'; 'I don't know her nearly well enough yet'; 'I need to go back and listen carefully to what he says again. I don't think I was picking up on all the nuances before'.

We return to one of the examples with which we began this Introduction: the student who wanted to carry out empirical research into whether being a Scout is conducive to the development of self-esteem. Behind this seemed to lie the assumption that we know what self-esteem is – that it is pretty much as unproblematic as the individual Scout's height or weight – the question being how to measure it and establish the correlation with Scouting. Now there are, in fact, other puzzling questions to ask about self-esteem, which require approaches more typical of *verstehen*. It is not at all clear that we know what 'self-esteem' *means*, that is – echoing the title of this book – how we are to understand it. Is it a good thing? Might some people have too much of it, so that we might think they would be better off with less? Would less self-esteem be the same as humility, and is humility a quality we would like to see more of in our Scouts, our children, friends, elected representatives in Parliament or Vice-Chancellors? Is it the same as diffidence, or is that something else? What are we to make of people whose apparent self-esteem – the air that they can take on anyone and beat them, perhaps – we suspect of concealing deep insecurity (in the UK we might offer as an example a