

## Sketch for a 'Demystifying' Approach to Social Research.

To begin with, I found the 'statistical responsibility (SR)' notion "progressive". It promised to make us reflect on at least three standard criticisms of an accomplished piece of research, viz:

- (i) it hasn't used significance tests (which it should have done) - or else it shouldn't have used them;
- (ii) it could have used controlled experimentation; and/or
- (iii) it didn't control for variable X.

Now statisticians are appropriately prepared to make criticisms (i) and (ii) since they refer to probability, perhaps the central concept in our approach; in the case of (iii), we are not so trained, unless we have studied the theoretical frameworks of the relevant substantive field. These frameworks specify which 'third variables' are plausibly relevant, and hence need to be controlled. Without a commitment to some theoretical framework, any conceivable third variable could be relevant, and criticisms of type (iii) could always be made (in particular by a statistician). If this is so, then whether a particular version of criticism (iii) is actually made, becomes a sociological, and not just a methodological, question.

Thus the notion of statistical responsibility is useful, to the extent that it clarifies the grounds -and the occasions-on which statistical criticisms are made.

However, SR is also limited. Consider two opposite extremes of research report: one which appears methodologically and statistically acceptable, but which purports to show that progressive methods of teaching are less "effective" than traditional ones (N. Bennett's research (1976) might seem to fall into this category at first sight); and the other which is methodologically flawed, but reaches a conclusion which is "progressive" with respect to previous research, for example, as in Rutter, that schools do make a difference to childrens learning.

Now the SR approach, if that is our only criterion for criticism, would seem to suggest that we should stop with the methodological and statistical points, without evaluating the research's political implications, and without critiquing them - and struggling for or against their implementation - appropriately. And that is why SR is limited. (For a much fuller discussion of related issues, see the final chapter of Demystifying Social Statistics).

SR is also limited in terms of the audience to which its criticisms are addressed: statisticians and other professional researchers and research consumers. The notion of 'demystification' orients us to a wider audience - our political allies with whom we can struggle to control the implementation of policies which are allegedly backed by any particular research findings.

To 'demystify' research output is to attempt to strip away the mystique that derives from it being 'research', based on 'statistics', perhaps produced by 'computer'. Against this we emphasise that research results are a 'social product'; that is, they are not, in any ultimate sense 'objective', or based merely on 'technical' decisions (as one might think if they believed that results were based only on data, manipulated mathematically). So we point to the role of commitments - theoretical and political - in providing the bases which structure both the data produced and the techniques used to analyse them.

In considering output from research as a social product, it may be useful to distinguish (i) research results, the "statistical" data and analysis which are themselves structured by commitments as above, and produced by conceptual and technical tools (see Ch. 9 of Demystifying); (ii) theoretical conclusions ('findings') where the results are interpreted within some more or less coherent framework of theoretical commitments, and (iii) policy conclusions ('recommendations') where the results and findings are put forward to justify some course of action.

(I intend this separation as an analytical - and very tentative - one, difficult to sustain in practice. However, as the three products are often associated with 3 different groups or institutions, viz. statisticians, disciplinary researchers, and "policy-makers", the separation may have some usefulness.)

So what does demystification involve? It includes:

1. illuminating what the results and findings are claiming on their own terms;
2. clarifying in whose interests/political commitments the findings and recommendations (are likely to) operate;
3. showing how various structures - social, organisational, and ideological - operate so as to mystify the social products (i), (ii), (iii) above), and
4. assessing the possibilities for re-use of the products within frameworks involving different commitments.

Let us look at how these four facets of demystification work out in connection with three products in the case of the Rutter et.al research.

#### 1. Illuminating results and findings:

Here we need to pay attention not only to the statistical tools - how they have been used, what are the required assumptions, and their limitations (as outlined in an earlier section of this report) - but also to the conceptualisation and operationalisation underlying the results and findings.

For instance, take the notion of school process which is central to the Rutter research, as that aspect of schools, rather than resources or physical facilities, say, which is most important in shaping pupils' attainment. Yet, school process does not appear to be particularly coherently conceptualised, especially when we look at the list of 46 facets of process (one with as many as 10 indicators) in Appendix E. The 46 variables are grouped into 9 classes (e.g. academic emphasis, teacher action in lessons, punishment, peer group stability, etc.), but we are given no clear idea as to how many dimensions the school process variable is thought to have, why the weightings used for various indicators are thought appropriate, or how the concept of school process is related to concepts developed in previous pieces of educational research. (We pay attention to the historical development of concepts, because it is only in relation to the conditions under which they are produced, and the concepts used previously, that we can fully grasp them. Hence to grasp the idea of demystification, advanced here, we need to understand that of statistical responsibility, and why the Education Group elaborated it, as well as uses of 'demystification' elsewhere).

So, the lack of coherence in the conceptualisation of school process alerts us to be skeptical of interpretations of the results which presuppose that school process is one dimensional, or which may be sensitive to variations in the weighting of various facets of process. (We could of course investigate the sensitivity of results to variation in either type of assumption, or to other assumptions).

Later in the report, the notion of 'school ethos' is brought in, as if from nowhere, seemingly to substitute for school process. At its first introduction (pp 55-56), it is characterized as the "climate of expectations and modes of behaving", but then, almost immediately afterwards, we are told that the school process variables concentrate on "happenings and behaviours", rather than "general attitudes which may be behind them". There may well be a confusion here between (professed) norms and (actual) practices - a distinction which we know to be important theoretically in the social sciences, and which could have important policy implications.

The researchers' reason for inferring the existence of an ethos is that "the combined effect (of school process measures) was much more powerful than that of any individual factor on its own". However, recall that the researchers produced a purified index of school process by selecting from the 46 facets of school process, the 39 which had a statistically significant correlation with any one of the dependent variables (listed earlier). Therefore, they have themselves produced the more powerful combined effect by throwing out from the combined index those other 7 facets.

The overall conceptualisation behind any piece of research is also worth examining, both for what concepts are included and what are not ('silences'). In this research it is clear that 'resources' are not included.

In the conclusions (No.5 p.178), we are told that differences in outcome were not due to "physical factors" measured such as size of the school, age of the buildings or the space available. By the next page, it is claimed that the main source of variation between schools in their effects on children does not lie in "factors like buildings or resources" (p. 179; my emphasis). From then on, it is assumed that it has been shown that resources don't matter - though there has been no attempt to measure them.

And yet, of the various facets, and corresponding indicators, of school process/'school ethos' in Appendix E, at least one quarter seem to be dependent on resource allocation to the school (e.g. library use, percentage of general (as opposed to specialist) teachers, percent teachers' time spent on assembling "equipment", outings, pupil conditions, decorations of classroom, clerical help, and so on).

On their own terms, these researchers have declined to bring resources into their conceptualisation of educational attainment and factors affecting it; hence no one could expect the research to allow study of the effect of resources. And yet at the same time, it seems to me that they have brought resources in, via their operationalisation of various facets of school process - but still they deny the importance of resources. Why should this be?

There are several ways we can attempt to answer this sort of question. First, we may have access, directly or through reports, to things said (or written), often not "in public", by the researchers, or by officials of the sponsoring institution, during or after the research. Or the researchers may publish a 'reflexive account,' describing the theoretical and political commitments underpinning the research, and the practical problems experienced in its execution, either with the final report, or later (in a book such as Bell and Newby eds.(1977), Doing Sociological Research). Or, lacking such immediate evidence, we may have to try to 'read' the political commitments in any formulations the researchers or others make of policy recommendations. I now turn to these.

## 2. Clarifying the interests behind findings and recommendations:

In whose interests is the emphasis on school ethos, at the expense of resources?

First of all, if accepted, these conclusions would reduce pressure on the providers of educational resources - LEAs (and rate-payers), the DES and national government.

The implications, in policy terms, of the claimed importance of school ethos are, I think, more subtle. Fortunately, the researchers' views of these are spelled out in some detail (pp.192 ff.). One quote will hopefully give some indication:

"The atmosphere of any particular school will be greatly influenced by the degree to which it functions as a coherent whole, with agreed ways of doing things which are consistent throughout the school and which have the general support of all staff". p.192).

One striking feature of the quote above is its ambivalence: it might be interpreted to suggest to teachers (1) that they allow plenty of time for discussions (formal and informal) with colleagues to shape rules and procedures and to evolve sensitive ways of dealing with conflicts, either with each other or with the pupils (a 'democratic' interpretation). Or the passage might be interpreted as suggesting (2) that schools need a firm hierarchy with some "guardian of ethos" at the top - who might be the head, or perhaps member(s) of the inspectorate, or - if it's a national ethos that were after - even an official in the DES! (an 'hierarchical' interpretation).

At the stage of interpreting the theoretical findings into policy recommendations, other groups besides the researchers get into the act. And to struggle to promote one policy interpretation rather than another may depend on factors (e.g. access to various media, use of prestige institutions or individuals), which leave issues of technical or theoretical competence behind. At this point, we can't afford to stay at home with our notions of statistical responsibility (or of theoretical coherence) - we need to be involved in the field of social activity where the struggle over policy is taking place.

There we become aware of the various competing interpretations of the research which will count practically. It appears that there may already be an official 'line' on the Rutter research. For example, a teacher in a London comprehensive tells of the Deputy Head returning from a DES-sponsored conference which discussed the research; the conclusions which the Deputy related to the teachers were that school ethos, and hence performance, were better where:

- (a) pupil and teacher punctuality were checked on;
- (b) teachers assigned, and checked up on pupils' homework;
- (c) senior teachers "chivvied along" more junior teachers.

Clearly this sort of interpretation is of type (2) (hierarchical) above, rather than type (1) (democratic). Either interpretation (indeed any policy interpretation) goes beyond the statistical results, and the theoretical findings. This shows the need for a response to research such as Rutter's to go beyond the making of statistical and theoretical points.

### 3. Showing how various social structures mystify.

Here I shall focus on two ways in which the statistical results themselves are mystified (leaving aside further mystification of theoretical findings and policy recommendations). The first comes from ideological structures, the second from organisational ones.

The report of the Rutter et al. research is very technical: it makes great use of the log linear model, presents many tables of analysis, and provides a technical appendix on the log-linear model which probably would be comprehensible only to someone who already understood the model. The report will be difficult reading for almost all teachers, almost all members of the public, and any non-statistically-trained academic. (This is not to say that even statisticians will be able to determine off the top of their heads the effects, if any, one the results of some of the blemishes in the method e.g. switching from 7 categories of verbal reasoning to 3, part way through).

The technical difficulty of such research means that most people are unable to understand the results, and hence they are unable to criticise those results. In addition, this inability often makes them extremely diffident about criticising policy recommendations which seem to be based on quantitative material - despite the fact that they are as competent as anyone to respond critically to recommendations which affect their interests. As long as this society continues to produce generations of whom the vast majority are largely non-numerate and are diffident about their ability to criticise an argument with quantitative aspects, research such as Rutter's will tend to be subjected to less scrutiny than it would receive in a more numerate society.

Thus moves to promote numeracy amongst the population at large, especially among teachers, may have significant political effects. My experience as a teacher suggests that improving numeracy is usually accompanied by even more important increases in confidence, and readiness to criticise quantitative arguments. In this sphere, too, the 'social product' approach - that results, findings and recommendations are moulded by commitments of various types which must be scrutinised as much as the statistical techniques and data used in the process - can be a liberating insight.

In addition, the sensitivity of the research results and findings to some of the flaws detailed in the criticisms given earlier can only be investigated by re-analysis of the data. Here it depends on who 'owns' the data and whether they will allow other researchers access to it. In the case of the Rutter research, it is not yet clear whether the Inner London Education Authority will allow other researchers, or teachers' groups, to re-analyse the data.

The S.S.R.C. has a very commendable rule that the data from large-scale research funded by them should be deposited in the Survey Archive at Essex. However, some pass through the nets: Neville Bennett whose SSRC-funded Teacher Styles and Pupil Progress was published in 1976, had not until recently got round to depositing the data from his controversial report. (He has just done so).

#### 4. Assessing the possibilities for re-use.

Here I can only sketch some ideas for re-using the Rutter data to investigate two questions raised by this discussion: (i) the importance of financial resources in producing educational attainment and (ii) differences in effectiveness of democratic, as opposed to hierarchical, approaches to enhancing 'school process' or 'effects'. (Obviously, these possibilities for re-use depend on gaining practical access to the data).

In the case of (ii) since we have no indication that the data was produced with the distinction between democratic and hierarchical approaches in mind, it is extremely unlikely that we will find any basis, even in the detailed operationalisations of school process, for investigating any differences. With (i) the situation seems to be as bad; however, as we saw earlier, the researchers appear to have produced inadvertently what may be acceptable indicators for resources, at least for the amounts deployed for certain crucial uses within the school, if not for the total amount allocated to the school.

We would then need to develop our own re-conceptualisation of the factors affecting educational attainment, including resources (something which I cannot do here) in order to give an independent concept of resources, against which the indicators for resources available in the Rutter research could be assessed. We could then proceed to re-analyse the data, using the indicators for resources and 'ethos' as separate independent variables. (An indication of what I have in mind may be found in Jim Lindsey's article in the June 1974 Comparative Education Review, Special Issue on Education Attainment, where he uses Marxist social class categories to re-analyse IEA data which had used census-type occupational categories as an independent variable).

To summarise, this section began with an acknowledgement of the strengths of the 'statistical responsibility' approach, but also with dissatisfaction over its limitations; some of these are outlined briefly here. In response, I develop the notion of 'demystification' of research reports, and illustrate four stages of this approach with respect to the Rutter research. This is still very sketchy and requires much more work: responses to the development so far would be most welcome.

Jeff Evans.  
Middlesex Polytechnic.  
Enfield.