Explanation and quantification in educational research: the arguments of critical and scientific realism

Roy Nash

* Massey University College of Education, New Zealand

Online Publication Date: 01 April 2005

To cite this Article Nash, Roy(2005)'Explanation and quantification in educational research: the arguments of critical and scientific realism',British Educational Research Journal,31:2,185 — 204

To link to this Article: DOI: 10.1080/0141192052000340206

URL: http://dx.doi.org/10.1080/0141192052000340206
Explanation and quantification in educational research: the arguments of critical and scientific realism

Roy Nash*
Massey University College of Education, New Zealand

(Received 23 July 2003; resubmitted 21 October 2003 and 28 January 2004; accepted 16 February 2004)

The influence of Bhaskar's critical realism has become increasingly evident in the debate on the nature of educational research. This philosophy, dedicated to the overthrow of positivist doctrines, in particular those of empiricism, nominalism, and causation as constant conjunction, is the foundation of contemporary proposals for new perspectives and possibilities in educational research. There are, however, aspects of critical realism, as it is presented, that may repay further reflection. This article suggests that the rejection of statistical modelling by critical realism is not supported by scientific realism. The argument is illustrated by substantive analyses of the relationship between poverty and educational attainment (reading) using the PISA (Programme for International Student Assessment) 2000 UK data set. It is concluded that an approach to the construction of explanatory narratives based on scientific realism is more likely to effect a principled integration of the theory and practice of qualitative and quantitative research.

Introduction

The debate over ‘qualitative’ and ‘quantitative’ research methods, particularly in as much as this has to do with the possibility of statistical modelling and forms of explanation, may have gained a new source of energy from the contemporary influence of critical realism on social theory (Bhaskar, 1993). Two recent texts, Scott (2000) and Willmott (2002), may suffice to show that this influence has been extended to the discussion of theory and method in educational research. These works give additional support to Manicas's (1998) critique—reproduced in a standard reader on critical realism—of the positivist assumptions of quantitative methodology and causal explanations based on regression analysis, which also takes

*Massey University College of Education, Hokowhitu Campus, Centennial Drive, Palmerston North, New Zealand. Email: r.nash@xtra.co.nz

ISSN 0141-1926 (print)/ISSN 1469-3518 (online)/05/020185-20
© 2005 British Educational Research Association
DOI: 10.1080/0141192052000340206
its examples from educational research. The semantic opposition of ‘quantitative’ and ‘qualitative’ is a self-evident mistake, as many have pointed out, but the institutionalised methodological division of research approaches legitimated by these terms is deeply rooted in philosophical theory, and it is in this context that the critique joined by critical realists merits some attention. The practice of critique is difficult and whether it is best directed at an ideal-type reconstruction or at a concrete text is often a nice point. The problem with the former is the likelihood that no one will admit to holding the position described, while the problem with the latter is that the text is easily declared to be unrepresentative. It will become clear that this article has adopted, in due acknowledgement of the risks, the second strategy. There is something very important and practical at stake in all this: what information do we need to have in order to explain the events, processes, and states of affairs educational research is concerned with; how can that information be obtained; and what form must explanations take in order to be useful?

The structure of the article may be found complex: it will argue from the standpoint of scientific realism that the rejection of quantitative modelling by critical realists, taking Scott as a leading representative of this position in education, is unsound. This project will require a brief comparison of the contrasted realist positions in the philosophy of science, and a concrete illustration of the problems posed by non-realist forms of quantitative explanation that their shared critique of positivism attempts to resolve. Critical realism and scientific realism are closely related positions, but whereas many critical realists have concluded that statistical modelling must be rejected, scientific realists maintain that quantification is necessary to science and that the conventional epistemological assumptions of applied statistics can be surmounted. Scientific realists, for example, argue that the assumptions of positivism, often implicit in statistical modelling, are not inherent to quantification itself and that statistical techniques may be used to enhance realist explanatory narratives. The thesis presented here does not claim to be original in disputing these matters with critical realism from a position of scientific realism. Kemp and Holmwood (2003), for example, convincingly argue that regular events occurring in the social world provide an ontological basis for quantification and non-positivist explanations. And Byrne (1998) similarly concludes that the rejection of statistical methods by critical realism is an error incompatible with the complexity and chaos of the way real societies work. It seems important, however, that these issues should be raised in the specific context of educational research, and perhaps also in a context that will show why all forms of realism share serious reservations about the language of positivism adopted by conventional approaches to statistical explanation.

Realism about the things of the world is a common-sense doctrine and it might be expected that a realist approach to educational research would broadly accept the established approaches that guide the conduct of scientific research and its related techniques of quantification. This is, as has already been indicated, largely so in the case of scientific realism, but critical realism has rejected the standard models of statistical research and explanation as incompatible with its philosophy. One would
need to be bold to summarise the elements of critical realism—Collier’s (1994) discussion requires an entire book—but Bhaskar (1998) has provided a short introduction that may be taken as authoritative. There are, he maintains, three core elements: (i) the things of the world cannot be reduced to the domain of the empirical, (ii) the domain of the real is more extensive than the domain of the actual, and (iii) the nature of the world is stratified in such a way that although one kind of mechanism may be grounded in another, it cannot necessarily be reduced to the elements from which it has emerged. Critical realism, it may be noted, is also a moral philosophy: it argues for an ontological ethics based on human well-being, and rejects the fact–value distinction (Collier, 1999).

As a philosophical movement, critical realism asserts that the entities of the world possess causal powers by virtue of their existence, and that the explanation of events, states, and processes should be made with reference to the properties that confer such powers. The application of this philosophy to the social sciences, including sociology, psychology, and economics, is being worked out by several authors, many of whom are linked to the Centre for Critical Realism at the University of Warwick. Bhaskar (1979) argues for a form of naturalism, an approach to science that recognises the fundamental unity of the world, grounded in the specific and emergent properties of physical and social entities. The recognition of a single ontology at this level enables Bhaskar to insist that although the social world is an open system ‘characterized by the complete absence of laws and explanations conforming to the positivist canon’ (Bhaskar, 1998, p. xv), there is no obstacle to the identification of systematic patterns of a kind that will allow the possibility of empirical controls for the purposes of scientific enquiry. Moreover, although Bhaskar understands that the social world is constituted as a result of meaningful activity, he explicitly rejects the idea that sociology is limited to the investigation of social events and processes where the intentions of agents can be ascertained, or that sociological explanations must necessarily be made in the light of such knowledge. The subject matter of sociology, on the other hand, certainly includes the beliefs that have been generated within a society and which affect the activities of its members.

Critical realism is particularly conscious of the reflexive manner, the so-called double hermeneutic, in which social theory in its own right can have a dialectical influence on social practice. It is a wide concept of realism that takes ‘absence’ as a definitive case of the real. As Bhaskar’s critical realism has developed it has made increasing use of the concept of dialectic to express the complex nature of entities and the relationships between them. Social life is thus constituted by material transactions with nature, by interpersonal relations, by the emergent social structures of such relations, and by the stratified constitution of the self. The explanation of social events, processes, and states of affairs requires an account of the mechanisms by which they have come to into being.

Bhaskar does not attempt to distinguish his critical realism from scientific realism, particularly in so far as realism establishes ontology as the necessary foundation for the practice of physical science, and all the indications are that he regards his position as a development within scientific realism. There may be no need to detail
the points of difference between Bhaskar’s critical realism and what might be called the standard doctrine of scientific realism, but it may be useful to indicate where the most significant deviations of critical realism from the general position are to be found. Bunge (1998) is among the most assertive contemporary advocates of scientific realism, and although in some respects idiosyncratic, there is no realist philosopher of science more systematic, and his position may be taken for this purpose as a benchmark. Scientific realism cannot be described as a movement, as Bhaskar’s critical realism has become, and in nominating Bunge as a representative there is no intention to overlook the contributions of other realists in this field, some of whom have already been noted, but it will be consistent with the approach taken here to focus on the stance develop by this writer.

Bunge’s philosophy is no more readily summarised than Bhaskar’s, but the important differences, in the context of this discussion, have to do with Bhaskar’s introduction of the concept of dialectic (which Bunge dismisses as worthless); an apparent support for explanations in terms of powers (where Bunge prefers properties); a tendency to argue that an effect demonstrates not only the existence of a causal power but the nature of the entity with that power (which Bunge regards as an inadequate demonstration); and a noticeable rejection of explanations supported by quantitative methods (which Bunge regards as essential to a mature science). There is, for want of a better term, an unmistakable tough-mindedness to Bunge’s philosophy of science apparent, for example, in his uncompromising materialism—he specifically rejects Popper’s World 3 (Bunge, 1981) which Archer (1996) has incorporated within her realist sociology—and, perhaps even more obviously in his intolerance of anything he considers to be non-scientific, including hermeneutics, psychoanalysis, and postmodernist treatments of ‘discourse’, all of which approaches, in the case of some writers more than others, have actually been given sympathetic treatment by critical realists. Scott, for example, treats ‘discursive formations’ as real mechanisms. Although Bunge’s blend of moderate empiricism, bluntness of expression, intolerance for theories some critical realists are willing to indulge (psycho-analysis, hermeneutics, dialectics, and so on), and preference for mathematical expression thus makes it unlikely that many critical realists will give it much attention, his work may be regarded as closer to the recognised stance of scientific realism.

There should be no doubt that the critique of statistical thinking has the authoritative endorsement of leading advocates within the critical realist camp. Writers accepted as critical realists have provided a trenchant critique of quantification in social science, dismissive in its implications, that is likely to be regarded as the definitive position of realism in its most general sense. Yet it would be a retrograde step should the view of quantification held by many writers influenced by critical realism become identified as the definitive realist position, for scientific realism has a very different stance on quantification. Bunge argues, for example, that the social sciences need more, not less, quantification, and that of a more rigorous and precise kind. This final break gives the current article its theme. It is proper to mention in this context, however, that Bunge, in these respects not
noticeably idiosyncratic, does not accept the conventional concept of measurement in social science, recognises the concept of emergent social entities, rejects the Humean concept of causality and, although maintaining a strong preference for scientific explanations based on established laws, finds the classical deductive-nomological model of explanation inadequate. This account has focused on certain manifest differences between critical and scientific realism, as these respective positions are developed by Bhaskar and Bunge, but it should go without saying that both share the common ontological ground of realism: they do not dispute that the entities of the social world are constituted by social relations, that these emergent entities have properties in accordance with their being, and that the explanation of social events, processes, and states of affairs requires an account of the real mechanisms that bring them about.

It is not altogether surprising, in the light of this commentary, to find that Scott and Willmott, drawing on critical realism, have developed themes at once hostile to quantitative analysis—and in Scott’s case at least—openly sympathetic to certain postmodernist influences. Scott’s work will be given close attention in this discussion as a text more systematic in its argument and located in an influential social research publication series. As I have made clear, to respond to a specific author and text is often an appropriate way to deal with the objection that positions subjected to critique have no substantive existence or influence. Scott is an eminent figure in curriculum studies and an active participant in the wider British debate on educational research. Whether or not the interest in critical realism will continue to gain ground remains to be seen, and there is no need to overestimate the current impact of this philosophical tendency on mainstream educational research, but it might be thought that Scott’s attack on quantitative models in this field is worth examination in its own right. There are serious matters to discuss that go right to the heart of what it is to conduct an effective scientific investigation and to construct an adequate scientific explanation. This account, as should already be evident, must also engage with the detailed arguments of statistical analysis. For this latter purpose the UK PISA (Programme for International Student Assessment) 2000 sample of about 9000 15-year-old students provides a resource that may give the illustrations—of the interrelationships between social class, wealth, literate resources, and attainment that constitute Scott’s own examples—some substantive value in their own right.

Scott’s critical realist objections to statistical modelling

Scott mobilises three arguments against mathematical modelling. There are, as he recognises, several approaches to mathematical modelling, but the critique is broad-based and distinctions between them are regarded as less significant than their common epistemological foundations. The first argument has to do with the characteristics of open and closed systems; the second with the distinction between association and causation; and the third with the dimensions of intensionality and extentionality in the social world. Scott argues that the open character of the social
world, which follows from the meaningfulness of social action, ensures that the associations reported by statistical analysis cannot be interpreted as revealing determined causal relationships. A systematic response to these arguments might be expected, but this is not necessarily the most effective procedure to adopt, and two of the arguments are rather easily dealt with.

The fact that the social world is an open system does not rule out the potential usefulness of quantitative models. Bhaskar (1998) accepts this point, in a discussion of Lawson’s (1997) use of such models in his own critical realist approach to economics, and there seems little more to add in this context. A similarly brief response can be given to the argument that that intensionality—the dimension of meaning—makes explanations based on quantitative modelling inappropriate. Bhaskar acknowledges that social science must attempt to explain social events, processes, and states of affairs regardless of whether the intentions of action, that is to say the meanings actors give to their actions, are known. To insist otherwise would unnecessarily restrict the scope of social science, for the intentions and interpretations of agents are not always available and, even then, cannot be taken at face value. The problems raised by the concepts of association and causation are, however, much more serious. Scott’s critique is basically sound in this area, but the matter requires a more extended discussion, and the implications are open to debate. In some respects, moreover, Scott’s objections are actually not the most telling that can be brought against quantitative modelling. The short answer to his case is simply that explanations supported by mathematical models do not need to make many of the assumptions he believes are necessary and, so the paper will argue, can be accommodated within scientific realism. This response may be elaborated through an analysis of a key example he gives of poverty and its effects on examination performance.

Scott argues that the association between poverty and educational attainment should not be regarded as one that reveals a causal relationship. In an open system, he asserts, a low income does not determine level of attainment because, among other things, the meaning given to poverty is not identical for different people and groups and, consequently, the association cannot be regarded as causal. As he writes (p. 44):

"The measures which are usually used to determine the relationship between poverty and achievement in school are approximations or proxies to the real variables which are implicated in the causal mechanisms which may or may not be operationalized."

This statement requires a little decoding. The term ‘real variables’ must refer to the properties of real social entities, and ‘implicated in the causal mechanism’ must refer to components of the causal mechanism, where ‘mechanism’ is identified here with a process that may or may not be activated. The observation that causal mechanisms may or may not be operationalised conveys the idea that a mechanism does not work unless, as it were, it is first switched on. In this context, it might be said that the presence of books in a home indicates certain literate capabilities of parents that may be utilised—but then again may not be—to instruct children in the arts of reading.

Scott develops his theme by arguing that, amongst other possibilities, poverty might cause poor examination performance due to lack of space, which means that
children cannot do homework; poorer diets, which mean that children cannot concentrate and so learn less quickly; low levels of cultural capital, which means that children cannot access ‘pedagogical interactions’; and low levels of parental emphasis on the importance of education, which means that children do not try as hard as they could. How quantitative evidence might be analysed within a realist framework will be examined for particular examples with reference to the PISA data.

Scott argues that poverty is a proxy for the real variables, by which it is clear that he means not an indicator variable of any kind, but a specific lived process. The argument here, by distinction, is not simply that an indicator variable may be a proxy for another, as a proxy for household income might be constructed from a selected list of household items (which is exactly what PISA does), but that poverty as a state of affairs is a proxy for an effective process, or mechanism, at a specific organisational site of the kind necessary to the construction of a complete explanation. This failure to distinguish between an indicator variable and the processes it points to, a confusion of model and reality, is one for which realists have no excuse. The constraints of thought imposed by positivist concepts, however, are so powerful that they can limit even those who possess the theoretical key by which their bonds may be released. Scott’s insistence on process as mechanism is sound in as much as, other things being equal, explanations that contain such information are to be preferred to those that do not, but it will be argued on realist grounds that this position should not be accepted without reservation and that it can have unwelcome implications for sociology.

The aim of sociology is to provide explanations of social events and processes. A complete explanation will include an account of social structures (properties of organisations), individual dispositions to act, and established practices (Nash, 2003). The best accounts of social processes are thus multilayered, and in their attempt to reflect the complexity of the world offer narratives that integrate rather than disintegrate, as is typically the case with reductionist models, many of which actually threaten to eliminate the structural level entirely (Mayer, 1997). The arguments necessary to examine the relationship between levels of explanation are highly complex and there is no suggestion here that the search for an effective process mechanism is inadequate in virtue of its reductionist character. When a kindergarten head teacher sees poverty as a barrier to learning because children arrive at school ‘hungry, badly clothed, unwell, unhealthy’ (Groser, 2003), she holds a position realists may support and endorse. Poverty is the state of being poor and as such it may be the cause of some other social state including relative educational attainment. At the level of the family, the state of being poor is a cause of low attainment whenever the lack of money directly or indirectly affects the capacity of children from poor families to learn at school. It is not necessary to know exactly how this happens in order to demonstrate the fact that it does.

Let us suppose that one of the conditions created by a low income is that children are left in poor-quality childcare with the result that cognitive development is inhibited. This will not be true of all poor families, but that does not mean that it is poor-quality childcare rather than the effects of poverty that is the cause of
underachievement with that origin (although it might well be useful to know that information). What we are dealing with is a hierarchy of linked causal processes. It would be pointless to construct explanations in terms of structural properties that exclude process mechanisms at the level of disposition and practice when these are known. But by the same token it is arguably no less satisfactory to reject the causal nature of structural properties when these are responsible for the generation of the effective dispositions and practices involved at those levels. Scott’s argument actually leads to a kind of reductionism that does not reflect the realist intuition of the sedimented character of social reality and is for that reason inadequate. There should, of course, always be an argument relating the variables and properties they represent in an appropriate way, but there are times when that mechanism must be postulated rather than demonstrated. A good case, for example, can be made for the argument that low-quality childcare is causally more likely to be experienced by low-income families, for it is all many can afford. These relationships, of course, are ones that can be tested most satisfactorily only by quantitative research, and the implications of that must be faced.

The argument that poverty can be regarded as a real state of affairs with causal properties may be compared with the position taken by Jackson and Pettit (1992, p. 119), who distinguish, in an argument entirely compatible with the principles of scientific realism, between process and programme explanations. In their words:

> The process explanation relative to any given level identifies the actual causes and relevant causal properties. The program explanation identifies a condition such that its realization is enough to ensure that there will be causes to produce the event explained: if not the actual causes, then some others.

Jackson and Pettit may be unwilling to regard poverty as a causal property, but they would surely recognise it as a condition that will ensure that directly causal events do occur. These authors note that ‘structural explanations explain, when they explain, by introducing factors which program for the realisation of the conditions explained’ (p.117). Poverty is a structural condition, and may be included in an explanatory model precisely because it is a condition that is likely to give rise to processes that will generate effective practices of the kind required to bring about the observed events. All of this suggests that a critical realist may argue that poverty is a state of being to which causal powers can be attributed.

Jackson and Pettit, for example, argue that high unemployment means that an increase in the number of working-class youths short of money makes it more likely that events will occur that will raise the crime rate, and this is also to argue that poverty—shortage of money—is a causal element in the whole process. But this is not quite the same as to argue, as Scott does, that poverty is a proxy variable standing for something else. This position, moreover, is not reductionist, as Scott’s is, because it recognises that poverty is a state of affairs that has whatever powers it has by virtue of the fact that it programmes for certain dispositions and generates certain practices. Pettit (1993) is inclined to argue that when a process explanation has been obtained a programme explanation is redundant, but the realist notion of a hierarchy of social states and entities is not incompatible with this position. There is
no need to accept that states of affairs, which are emergent at the level of the programme, do not have the causal powers attributed to them. This will be an appropriate point at which to explore these ideas with substantive illustrations, based on the PISA UK data, of what can be learned from a quantitative analysis considered from the position of scientific realism.

As the following sections present data from the PISA study it will be necessary to provide basic information on the sample and variables included in the analysis.

**The PISA sample and some core variables**

The UK Programme for International Student Assessment (PISA) data set provides information on about 9000 15-year-old students. The data can be downloaded free together with all necessary documentation and command files from the official website (Organisation for Economic Cooperation and Development [OECD], 2001). The study is a characteristic product of the international educational research community and is principally designed to facilitate international comparisons of educational standards. All the data are obtained by tests and questionnaires completed by principals, teachers and students. The analyses reported in this analysis use attainment in English and certain data from the student questionnaire. We are interested in socio-economic status, cultural capital, and poverty: the first task is to define the variables.

- **Reading Attainment**: Several estimates of reading attainment are provided by PISA. The analysis reported here uses the Warm estimate of reading attainment, which is the most useful for within-country analyses and recommended as the measure appropriate for reporting individual attainment. The estimate has a nominal international mean of 500 (UK, 522.6) and a standard deviation (SD) of 100 (UK, 100.3).
- **Socio-economic status**: The PISA SES indicator is based on the International Socio-economic Scale (Ganzeboom et al., 1992). Information about parental occupations is obtained from the student questionnaire, and the analysis given here uses the highest reported occupation of a student’s parents. Index values have been grouped into quintiles for the purposes of Table 1. SES 1 designates those in the upper fifth of the distribution.
- **Number of books in the home**: The student questionnaire asks students to estimate the total number of books in the home. The item informs students that there are about 40 books per metre of shelves and that magazines are not to be included.

| Table 1. Correlations between reading, wealth, SES and number of books |
|--------------------------|--------------------------|--------------------------|
| Reading                  | Wealth                   | SES                      | Number of books |
|                         | .109                     | .285                     | .365            |
| Wealth                  | .350                     | .232                     | .350            |
| SES                     |                          |                          |                 |
The categories provided are: none; 1–10; 11–50; 51–100; 101–250; 251–500; more than 500. The analysis presented in Table 1 divides the scale at the point closest to the median, resulting in groups with <100 books and >100 books. This is adopted as an indicator of ‘cultural capital’, although other indicators could be used, and a scale intended for this purpose, constructed of a number of items about family practices in relation to cultural consumption, is included in the PISA data set. The number of books in the home, however, has the advantage of transparency. As far as the substantive analyses presented in this article are concerned, it makes very little difference what index is used.

- **Wealth:** PISA provides a scale based on household goods to indicate family wealth. Students are asked to report the presence in their home of a dishwasher, a room of one’s own, educational software, an Internet connection, and number of cell phones, televisions, computers, motor vehicles, and bathrooms.

The statistical relationships between these variables will be explored in substantive analyses presented for the purpose of illustrating the core theoretical arguments in the discussion of realism and its critiques. Of course, it is not to be taken for granted that anything much can be learned about the relationship between family resources of wealth and literacy and educational attainment from the PISA data. The data are obtained from student questionnaire responses, and the degree to which these can be accepted as a trustworthy source of information of actual states of affairs is open to question. The ‘measures’ used, which are much better referred to as indicators, are also accepted in the form given by PISA analysts, and this involves a further set of assumptions. There are, indeed, to adopt a conventional term, likely to be serious measurement problems. The SES scale, in particular, is less than ideal, but then it is a great deal better than nothing. These matters are not, however, the focus of the debate, which at this point is concerned to show that the quantification of relevant evidence is necessary to support hypotheses about the processes that constitute the link between structural properties, in this case poverty, and the properties of people and the social practices they adopt. As Kemp and Holmwood (p. 179) conclude, ‘While … statistical regularities cannot, by themselves, establish a causal link, it is also important to acknowledge that the search for such patterns can be an important part of the process of identifying causes operating in the social world’. The procedure advocated is fundamentally that of science itself, and the rewards of the comparative method (that is to say Mill’s method of difference) can often be achieved with no statistical technique more problematic than counting. It is also important, however, within this realist defence of quantification, to draw attention to the difficulties posed by the residues of positivist language in statistical modelling.

**A quantitative investigation of poverty, overcrowding and attainment**

The search for the generative mechanism, where that is a social process, must often require quantitative forms of investigation. Scott’s supposition that poor families have less room to live in and therefore find it harder to complete their homework with the consequence that they learn less and so gain lower achievements in
examinations, moreover, can be treated as a set of related hypotheses open to test. It should immediately be clear, however, that without the use of statistical methods there seems no way to test any such hypothesis. One would have to show, for example, that poor children living in overcrowded homes achieve less at school than poor children from homes with adequate space. It is significant that Scott does not discuss how his suggested mechanisms might be shown to exist by any method other than direct observation. It follows, therefore, that he is also unable to investigate the contribution different but related process mechanisms might make, but there is no reason why cramped living conditions, poor nutrition, lack of ‘cultural capital’, or parental press for education, should not all be involved in the generation of class differences associated with an index of poverty and to various degrees that could, in principle, be assessed.

One way to discover whether the causal processes suggested by Scott actually exist is to treat the suggestions as hypotheses and subject them to test. Consider, for example, the suggestion that poverty has its effects because low-income families often lack the space to provide children with a quiet place to study and that as a consequence their attainments are lower than would otherwise be the case. Some elements in this causal chain can be tested with the PISA data. If the hypothesis is correct then we should expect that children from low-income homes, in comparison with others, often lack a quiet space to study, spend less time on homework, and have lower attainments. The first thing to note is that only 13% of students in the lowest quintile of the Wealth distribution actually report that they have no quiet place to study. The hypothesis is already looking a little shaky. The next step is to look at the differences between poor students with a quiet place to study and those without, and it turns out that they are relatively minor. The proportion of poor students reporting that they have no time for homework, with and without a quiet place to study is 7% and 13%. In the remaining categories the figures are, respectively: less than one hour, 33% and 28%; 1–3 hours, 53% and 45%; and more than 3 hours, 10% and 12%. The most one can say about this is that the trend is in the expected direction, but it is hardly convincing support for the hypothesis.

The final step is to examine whether the mean reading attainments of these students are actually different. It so happens that, with the exception of those who report having ‘no time’ for homework (in which case it would seem not to matter whether they have a quiet space in which to do it or not), the differences are not significant. We find that the means of those who do less than one hour of homework, with and without a quiet space, are 489 and 484: in the remaining categories the means are, 1–3 hours, 522 and 520; and more than 3 hours, 540 and 532. Scott argues that quantitative models provide no useful information about the causes of social differences in educational attainment, but the causal mechanisms he suggests as generating the association between poverty and low attainment cannot be observed or tested without reports confirmed by statistical analysis. If one is prepared to accept the self-reports of 15-year-old students as observations (which of a kind they are), then those analysed here give the hypothesis very little support. The difference in educational attainment associated with poverty does not seem to be
caused by this particular mechanism. There must be some mechanism involved, to be sure, but they will not be detected without exactly this kind of scientific procedure. It is incoherent to argue that a social process might be involved in the generation of the effect to be explained and yet to deny the value of quantification in testing the validity of that argument.

**Socio-economic status, cultural capital, wealth and reading**

The object of this section is to give a practical illustration of how the approach to quantification supported by realism might enable multilevel explanatory narratives, as opposed to multivariate causal models, to be constructed. This will involve the rejection of positivist concepts but the acceptance of what can be learned from statistical models. If scientific realism is to incorporate the findings of quantitative models then it must do so within an explicit epistemological framework. The investigation may begin with an examination of the pattern of association between the indicator variables. The PISA variables presented earlier are related, as might be expected, but the level of association is quite moderate. Table 1 reports the correlations. It will be observed that the correlations of Reading with SES and Number of Books are similar, and somewhat lower between Reading and Wealth. It is worth noting also that the correlations between SES, Wealth and Number of Books, arguably an indicator of ‘cultural capital’, are no more than moderate. Even the correlation between SES and Wealth—between variables that are often treated as equivalent ‘measures’—is not strong, although it is one of the largest in the set. This pattern raises questions that will receive further comment. It may be added, in this context, that the relationship between these indicators and what they point to is precisely the concern of the entire discussion of this article. With respect to the purposes of the illustration, it does not matter whether, for example, Number of Books is a good or a poor indicator of the property of families recognised as ‘cultural capital’, but it matters critically how that relationship is to be constructed in theory. If this remark seems obscure, it may become clearer as the argument is developed.

The multiple correlation coefficient between Reading and the three other variables, SES, Wealth, and Number of Books, attempts to express the total variance in Reading scores accounted for by those variables. The multiple correlation coefficient, calculated by a stepwise procedure, shows that each variable has a certain independent contribution to make:

- .351: Number of Books
- .419: SES
- .420: Wealth.

Does it make sense to ask which of these variables has the greatest effect on attainment? It is a question that quantitative models seem to have the potential to answer, and it is certainly possible to obtain answers that seem more or less plausible. The pattern suggests, for example, that Wealth has little independent
effect on reading attainment when SES and Number of Books are taken into account. It is a little more difficult to ascertain the degree of independence between Number of Books and SES. When SES is entered into the equation first and Number of Books second, the multiple correlation steps are .336 and .419, and it is thus clear that these variables are related to reading attainment scores in similar respects. This interpretation is confirmed by the path analysis showing the relevant beta weights given in Figure 1. The model supposes that SES (parent occupation) and Wealth are family resources that enable the purchase of books, as resources of literate practice, and that the pattern of interaction between these variables can be used to support theories about families, their resources, and practices, and their consequent effects on reading attainment. The pattern suggests that income level, as such, is a less important resource than certain others presumably associated with occupation. To the extent that income has an effect on reading it does so mainly through the possession of books. This is not quite so true in the case of occupation. What all this means, however, for a narrative of explanation has still to be worked out.

The statistical models presented here can easily give the impression that the contribution of such variables, and thus the effects on attainment of the structures and practices of families they represent, are somewhat trivial. After all, the proportion of Reading score variance accounted for by these indices of family resources is just 17.6%. The pattern of associations, however, can appear in a much different light when presented in a tabular form (Boudon, 1974).

Table 2 shows the mean reading estimates for students categorised by SES, number of books in the home, and an indicator constructed from items included in the Wealth variable. The information is fundamental to the discussion and it will not be redundant to draw attention to some of the most salient points.

In every SES group students with 0–100 books have lower mean scores than those with 101–500+ books in all categories of wealth. The within-class mean attainment difference associated with number of books ranges from 31 in the case of low SES students to 55 points for high SES students. The within-class difference between wealthy and poor students is relatively minor and ranges from 2 points for high-quintile SES students to 13 points for those in the mid-quintile SES category. In the
high SES category the difference in mean attainment associated with number of books and wealth is greater than in the low SES category, but there is no trend in the intermediate SES categories: the differences, from SES 5 to SES 1 are: 34, 44, 42, 40 and 63. This pattern is consistent with the regression models. The effect of number of books apparent in Table 2, even on a dichotomised scale, can appear dramatic. The mean attainment of ‘rich’ SES 1 students with 0–100 books is not significantly different from the mean of ‘poor’ SES 4 students with 101–500+ books. Such information is not immediately apparent from an inspection of regression models.

The value of statistical models depends on the interpretation given to the indicator variables. In the standard, positivist, approach to quantitative analysis, indicator variables are regarded as measures of an underlying concept that is supposed to be assessed more or less adequately by the standards of ‘validity’ and ‘reliability’ that prevail in this field. This is not the perspective of realists engaged with quantitative analysis. Bunge (1998), for example, has criticised the conventional approach to the measurement of social variables as fundamentally flawed. The most systematic critique of measurement theory relevant to this discussion merits more attention from realists interested in these problems than seems to be the case. Berka’s (1983) materialist theory of quantification argues that whether or not a property can be quantified depends on its nature and thus establishes ontology rather than

<table>
<thead>
<tr>
<th>SES</th>
<th>Wealth</th>
<th>Number of books in home</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0–100</td>
<td>101–500+</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>101–500+</td>
<td>n</td>
</tr>
<tr>
<td>1. High</td>
<td>Rich</td>
<td>534</td>
<td>99</td>
<td>587</td>
</tr>
<tr>
<td></td>
<td>Inter.</td>
<td>545</td>
<td>258</td>
<td>598</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>524</td>
<td>72</td>
<td>606</td>
</tr>
<tr>
<td>2. High average</td>
<td>Rich</td>
<td>522</td>
<td>193</td>
<td>560</td>
</tr>
<tr>
<td></td>
<td>Inter.</td>
<td>526</td>
<td>454</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>526</td>
<td>118</td>
<td>583</td>
</tr>
<tr>
<td>3. Average</td>
<td>Rich</td>
<td>502</td>
<td>173</td>
<td>536</td>
</tr>
<tr>
<td></td>
<td>Inter.</td>
<td>513</td>
<td>607</td>
<td>549</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>494</td>
<td>242</td>
<td>539</td>
</tr>
<tr>
<td>4. Low average</td>
<td>Rich</td>
<td>488</td>
<td>87</td>
<td>539</td>
</tr>
<tr>
<td></td>
<td>Inter.</td>
<td>493</td>
<td>594</td>
<td>542</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>495</td>
<td>293</td>
<td>531</td>
</tr>
<tr>
<td>5. Low</td>
<td>Rich</td>
<td>482</td>
<td>89</td>
<td>501</td>
</tr>
<tr>
<td></td>
<td>Inter.</td>
<td>476</td>
<td>779</td>
<td>506</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>477</td>
<td>529</td>
<td>516</td>
</tr>
</tbody>
</table>

Note: SES is derived from Highest International Socio-economic Index and divided into quintiles. Reading attainment is the Warm estimate for overall reading performance. Number of books is divided roughly at the mean of the seven-point category scale. Rich students are defined as those in the upper quintile and poor students those who live in the lower quintile. Students falling between these extremes, in the three middle quintiles, are defined as intermediate in wealth.
epistemology as the foundation of measurement. Berka distinguishes between scaling and measurement proper, proposes that the object of measurement (such as length) and the measured object (such as a stick) should not be conflated, and shows that the concepts of validity and reliability, which carry the burden of the positivist approach to measurement, are inadequate. An indicator of occupations, basically ordered by education and income, reflects the class position of a family; the number of books gives some information about the level of literate resources possessed by a family; and a list of possessions shows that a family has enough disposable income to afford them. What frames of mind they indicate and what practices they support are matters to be settled by investigation and analysis. Such research, of course, must include precisely those forms of enquiry rooted in the qualitative tradition: it is becoming clear that statistical explanation is a misnomer as far as scientific realism is concerned.

Indicators of occupational status and number of books in the home, for example, should not be regarded as measures of underlying concepts, but as scales on which homes are ranked by occupations, more or less ordered by income and education, and the number of books they possess. The assessment of attainment—in the PISA data set not technically a score—is derived from test responses transformed to provide a standardised index of performance. All these indicators point to properties of social entities or individuals, and the business of interpreting the ‘meaning’ of indicators is actually that of providing an argument that links an indicator to the property it is intended to point to. The relationship between an indicator and the properties it points to, whether these are structural, dispositional, or behavioural, is a complex matter that, fortunately, need not be discussed extensively in the case of the illustrative models presented in this section. It can certainly be argued that an estimate of reading attainment, a pseudo-score, points to both a relative performance level and an effective disposition to demonstrate such a performance. Class location is itself given by a complex of social relations, definitively those to do with the control of capital and labour, and these are relations of power that enable certain practices to be adopted. An indicator of class position, therefore, is one that gains its power to achieve the effects detected by statistical analysis through a great variety of class-associated practices.

What interpretation can a scientific realist place on the statistical analyses presented in this section? The regression equation shows that an increase of 1 standard deviation in SES is associated with an increase of 0.34 standard deviations in Reading. Similarly, an increase of 1 standard deviation in Number of Books is associated with an increase of 0.37 in Reading. The statistical interpretation of this is straightforward, but the substantive interpretation imposes very severe difficulties. The scales are obviously not comparable and standardising them, which is a routine matter, does nothing to solve the real problem. What it means, in fact, is that having a father who is a mechanic and a father who is a teacher is associated to about the same degree with reading attainment as does living in a home with 70 rather than 400 books on the shelves, for these are the approximate ranges covered by one standard deviation on each scale. There is no technical solution to the problems
raised by this observation. We know that occupation and number of books have powers of this kind, but how they are actualised, that is through what dispositions and practices, and even at what sites, cannot be learned from this information. Even if the implied equivalence could be given an agreed interpretation, which is not the case, one would still be faced with the problem of constructing a narrative of causal processes. In fact, any substantive interpretation would have to be such a narrative. This is likely to be an area where reality is discontinuous, and subject to breaks with a qualitative effect, rather than linear in character.

Perhaps the most plausible interpretation one can place on the overall analysis made in this section of the relationship between reading, wealth, socio-economic status and number of books is that family practices associated with the possession of books do contribute to the development of certain cognitive and non-cognitive dispositions effective in generating differential reading performance. The explanation in its most complete scheme has a structure-disposition-practice form that requires the adequate description of structures (properties of social entities), dispositions (properties of individuals that lead to action), and practices (established ways of doing things and hence structures of agency). Indicators of social class, wealth or poverty, and the number of books reflect structural properties of families, these properties are causally associated with certain dispositions, and these are dispositions that lead to the adoption of certain practices. It is in this way that the number of books in a home is most likely to have a causal connection to the attainment of children.

Jackson and Pettit’s argument that a programme explanation is possible, where social class is recognised as a cause of educational attainment even though the chain of transmission is not clear, may be accepted as legitimate. An explanation having the full structure-disposition-practice form is, however, always to be preferred when it is available. Yet this is not to say, with Scott, that poverty (or wealth) is a proxy variable for unknown resources and for unobserved practices with causal efficiency. The search for mechanism, moreover, requires some philosophical care. Bhaskar’s argument that the extraction of surplus value is the hidden mechanism determining the nature of class relations and class-associated practices clearly does not locate the effective mechanism at the level of practice, as in this example. If the realist concept of ontological level is to be respected, at least of the individual and of the social, then accounts that attempt to recognise properties at the level of structures, dispositions, and practices are likely, in one respect or another, to represent the effective causal mechanisms sought to explain the events and processes under investigation.

The interpretation of these statistical models, based on regression and on tabulated analyses, points to the need to utilise such models as a basis for obtaining knowledge of the real social processes they represent. The models are not, in themselves, treated as explanations, but as sources of information more or less useful in the construction of complex explanatory narratives. These explanations, moreover, will often require information external to the statistical model that is likely to be generated by so-called qualitative studies. The intuition behind Scott’s critique, in all likelihood, is grounded in the fact that the correlations between SES, number
of books in the home, and attainment, although presented in the language of ‘variance explained’, provide substantive explanations only at the programme level and leave the processes involved, the specific actions of people with certain dispositions at given sites of practice that actually ‘make the difference’, as secret as they ever were. And when it is suspected, as critical realism does suspect, that the language of statistical positivism bears the taint of ideology, the accounts of statistical modelling seem not only inadequate and incomplete but systematically misleading and therefore to be rejected. This is not the more considered response of scientific realism, but it should be understood that the genuine problems raised by this grounded discussion of how statistical models might be interpreted in a realist framework emerge from a common rejection of the positivist epistemology embedded in the conventional discourse of statistical explanation (Nash, 2002).

Conclusion

Realism rejects the account of causation as constant conjunction, opposes the nominalist interpretation of variables, regards models as representations of reality, and argues that explanations should, wherever possible, be given in terms of generative mechanisms. This programme suggests that the identification of causal factors requires a conceptual analysis as well as a technical analysis. If it is the case that there are multiple effective causes, each contributing some determinate effect, then that is a reality that explanatory models should attempt to represent. This is the essence of a realist approach to science. It is a central, if not actually definitive, problem of the sociology of education to provide explanations of how differences in social access to education are generated. There will be no dissent to the proposition that this state of affairs has multiple causes. Scientific realism, therefore, must accept the need for quantitative analysis to establish both the extent of social disparities and the estimation of the relative weights that should be accorded to distinct processes, the evidence for which can be ascertained, in part, by the statistical analysis of indicator variables. The positivist notion that the number of books in a home and other literate resources (such as computers, musical instruments, and so on), can be treated as indicators of a ‘latent variable’ representing, perhaps, a ‘concept of cultural capital’ that can be constructed as an object of measurement, and represented as a causal factor in the context of a model, is dubious in any version of realism. That criticism, however, is not in itself sufficient basis for the rejection of the forms of statistical modelling, illustrated in this article by analyses of the PISA data set, as useful in the construction of realist explanatory narratives.

On the contrary, scientific realism, as articulated by Bunge (a professor of mathematical physics before his switch to the philosophy of science) and many others, is entirely compatible with quantified approaches to social science. The critical realist critique exemplified by Scott must, therefore, be considered faulty on several grounds. In summary, the objections to statistical modelling are unsound; the attempt to deny causal powers to poverty as a condition of being is inconsistent with realism; the argument lends support to forms of reductionism that would eliminate
sociological variables and, finally, the position reached seems entirely to disallow the scientific testing of hypotheses about the existence of the process mechanisms believed to be responsible for the differences these accounts attempt to explain. If the cliché will be allowed, one might say that the critical realist critique of quantification and statistical modelling risks throwing out the scientific baby with the positivist bathwater.

Realism constructs explanations of events, processes, and states of affairs in terms of the generative mechanisms that give rise to them. In social science, including sociology, history and economics, the most complete explanations will have a structure-disposition-practice scheme where the structural properties of emergent social entities, the dispositional properties of individuals, and the actions performed by individuals within recognised social practices, are all included in the explanatory narrative. In as much as social structures, so conceived, are actually a level of reality constituted by pre-existing social relations and the results of historic practices, explanations at that level cannot necessarily be reduced to individuals' states of being or to their actions. The implications of this realist conception of the social, its properties, their reflection in indicators, and the construction of statistical models, must be given careful consideration. There are, to be sure, some fundamental problems with the conventional discourse of statistical analysis; the legacy of positivism is so deeply entrenched in its technical vocabulary that the construction of realist explanatory narratives is much more difficult than would otherwise be the case.

Scott’s critique is, at best, only partially cogent and cannot provide an alternative approach to the explanation of social events, processes, and states of affairs. It fails to recognise that employment of the discourse of positivism does not necessarily denote a formal allegiance to the theoretical tenets of positivism. As Byrne, who might be described as an idiosyncratic scientific realist himself, puts it: ‘Positivism is dead’ (p. 37). Its bones, however, remain fossilised in the lexicon of applied statistics. The terms quantitative researchers are constrained to use appear to commit them to a Humean concept of cause, and to a view of science that rules as illegitimate any reference to a reality outside its models, but this is not the practical view of most researchers in this field (Patterson & Goldstein, 1991). The same point can be made in relation to correlation and causality.

All textbooks in statistical methods advise students of the distinction between correlation and cause, and usually discuss the relationship between statistical models and the reality they are designed to represent. Although it is true that these discussions are never satisfactory—hardly surprising given that the problems are impossible to resolve with the only concepts available—it does seem unjust to suppose that statistical researchers are necessarily committed to the formal implications of their theoretical language (Bartholomew, 1999). It is possible to maintain that the epistemological foundations of statistical modelling do not render its substantive analyses inherently worthless. The conventional theory, so well entrenched as to be regarded by many quantitative researchers as nothing more than ‘the language of statistics’, nevertheless continues to pose severe difficulties. It is worth being clear about what these are.
The first point is simply that even to distinguish between the model and what it represents is a problem. Scott is hardly alone in having difficulty in speaking of the object of investigation without recourse to the word ‘phenomena’, or in being driven to use the terms ‘variable’ and ‘factor’ to refer both to an indicator and to the property that it points to. The substantive causal process is actually referred to as a variable, although it is plainly not, in this discourse. But if the conventional position, which is empiricist and nominalist, prefers to talk about models, where realism prefers to talk about the entities and properties they are about, both face a similar difficulty in making the link between model and the reality it represents. The fact that realism is committed to the world and the possibility of its demonstration does not make the business of its demonstration, when that is a bit more difficult than kicking large rocks, any the easier.

It would certainly help were realist thinkers in this field willing to adopt a set of concepts that makes it possible to speak about indicators rather than ‘measures’; of events, processes and states of affairs rather than ‘phenomena’; of variables as a reference to terms employed in a model; and of properties when referring to those features of the social world that have provided the data expressed by a variable. These small reforms alone would bring greater clarity to the debate about the nature and conduct of educational research. The problem of linkage—which positivism was designed to bypass and which quantitative researchers therefore so struggle with—is impossible to resolve with the concepts of positivist theory. In these respects, then, critical and scientific realism are agreed, but in the argument of this article that critique supports the case not for the abandonment of statistical models and explanations based on them, but the adoption of a realist epistemology sufficiently robust to improve the multilevel explanatory narratives of a naturalised social science.

These problems have been recognised by scientific realism, which, as several writers have now pointed out, is able to provide a non-positivist framework for quantitative modelling. Such a framework must allow the structural properties of the social formation—poverty is one among many—to be included in a causal narrative at the appropriate level. This is not to deny, of course, that poverty, as a social state of affairs, has its effects on, for example, educational attainment, as a result of dispositions as states of mind, and through social practices generated in some form as an adaptive response to poverty. Some implications of these philosophical concerns have been explored with reference to statistical models using the PISA data. The primary purpose of these is to ground the discussion in concrete illustrations, but if they make any substantive contribution to the explanation of social differences in education that is a bonus not be despised.

References
Bhaskar, R. (1979) the possibility of naturalism: a philosophical critique of the contemporary human sciences (Hemel Hempstead, Harvester).