

X

## APPLYING RESEARCH TO EDUCATION.

Lawrence Stenhouse, Director, Centre for Applied Research in Education,  
University of East Anglia, Norwich, England.

As a starting point I shall define research as 'systematic inquiry made public'. Like all such definitions this is too simple. However, it alerts you to my point of view and puts research in a particular perspective, and I hope this will serve to relate my argument to your own position.

Inquiry is a teleological pattern of action whose purpose is satisfaction, and it is related psychologically to curiosity, a disposition to explore the environment in order to assess its potential for yielding satisfactions. Inquiry readily becomes systematic in human beings and perhaps in animals, though the film Jane Goodall and the Wild Chimpanzees documents rather well the limitations of a chimpanzee's inquiry into the nature of a mirror and the way it peters out to be succeeded by other objects of attention. Classically systematic human inquiry is structured upon 'understanding', which gives it potential for continuity and development.

I choose the word understanding because it has two opposites which link the personal and the social. When we oppose not understanding to understanding we imply that there is a psychological basis for making the claim: 'I understand', and a similar basis for the claim: 'I do not understand'. When we oppose misunderstanding to understanding, we imply a public assessment of that claim. To misunderstand is to have an unjustified experience of understanding. One cannot say 'I misunderstand': only retrospectively after understanding can one say 'I misunderstood'.

The public aspect of the systematic inquiry that we institutionalise as research is largely concerned with the discrimination between

understanding and misunderstanding. From the point of view of the individual this is a process of testing one's claim to understanding in the appropriate public forum, as I am now testing my understanding by submitting my work on applying research to education to members of B.E.R.A. From the point of view of a research tradition the discrimination between understanding and misunderstanding hinges on the building of publicly accessible interpretations of theories, which are susceptible of improvement. This implies public criteria for judging what constitutes an improvement - that is, metatheories - and behind these a study which regards the theories and metatheories as problematic. This is epistemology, the inquiry into the basis of claims to knowledge; and this itself has behind it the area of inquiry which Aristotle called metaphysics and is now that branch of metaphysics called ontology, whose inquiry into the nature of reality necessarily has implications for epistemology.

Within this publicly institutionalized pursuit of understanding the great achievement is scepticism, that is, the capacity to treat all knowledge as provisional and susceptible to revision. Without public support it would be impossible to live with such abstract uncertainties as modern knowledge offers us and we should all revert to common sense.

This public possession of the standards and criteria which underwrite the possibility of scepticism is one main reason why our definition of research must include the idea of making public, of publishing. A second main reason is that research is collaborative. It enables each researcher to use the work of others, to stand on many shoulders. Thus it is important that published work should be presented in forms which are at once accessible to criticism and utilizable by others as contributions to their own work.

This is the essential basis of Wissenschaft, which I shall translate as scholarship, since science has a restrictive meaning in English. What I have been saying is true of history and of literary criticism and of philosophy, not merely of natural sciences and of human sciences.

When I address the problem of the application of research to education, I conceive it in terms of research lodged within the broad tradition of scholarship which I have just sketched. And, of course, the crucial issue in education, as in other applied fields, is that of the relationship of scholarship and research to action.

Classically, the field of scholarship whose implications for education are clearest has been seen as philosophy. A long tradition stemming from Plato and Aristotle associates educational action with ethics and politics. Education is seen as a means of promoting the Good, and the central problem for scholarship and research is to determine the Good. Or to put it in more fashionable, though equally crude, terms, the major function of scholarship is to determine the aims of education, and convince the educational practitioner of the correctness of these aims. The responsibility of the teacher is to pursue the right aims. In Britain this view of the problem is rather out of fashion at the moment. But an Austrian gymnasium teacher who took the view that it was her job to teach what is right without being deflected by the desire to see that students learned successfully attracted some support and understanding opposition from continental participants.

For most of this century the tide has been flowing in other directions. First there was an emphasis on understanding the pupil as a child, which was associated with the child-study movement. This was, perhaps, a first reaction to the recognition that rather humble aims of education such as teaching people to read were difficult of achievement. It seemed that the teacher might make the error of treating children as if they were adults and consequently fail in the educational task with the majority who could not adjust to the strange ways of schools.

Second, there was the growth of faith in experimental psychology - particularly in learning theory - as a basis for improving educational action. The teacher should have some expertise in understanding the learning process.

Third, there was the impact of mental testing with the concomitant attempt to use measures of aptitude as a basis for differentiating the aims and the methods of education among the school population.

I cannot in the time available to me here review the applications of scholarship to education which have been attempted in this century. Suffice it is to say that a major trend has been towards a concern with the exploration of the conditions within which the educational activity takes place, and particularly the psychological, and later the sociological, conditions. This line of research was followed with great enthusiasm by scholars in universities, and I attribute at least some of that enthusiasm to their perception that here was a prospect that the findings of pure psychologists and pure sociologists could be applied to the practice of education. If this were true, then research money provided out of concern for the improvement of practice could be legitimately used by scholars whose interests - in every sense of the word - were best served by the pursuit of pure theory - aimed at understanding the nature of the situation - rather than applied research - aimed at improving practice. And this tradition continues. Government funds for educational research flow in good measure to those whose curiosity is psychological and sociological rather than educational. The funds are not of themselves ill-spent, but there is a danger in their recipients' claiming more immediate relevance to education than they can justify. When they do so, distorted applications are made. It is difficult to know whether more blame lies with the sociologists or the teachers when working-class children are treated as linguistically stupid!

I would go so far as to doubt whether most of the work in psychology and sociology which is taught to prospective teachers has any close relevance to educational practice. Its relevance is to educational theory, that is, to a kind of theory which is to be tested by educational practice. The error is to attempt to apply results to educational practice rather than to use educational practice to test results. And the error is made by the researchers in the first instance. They

appear to claim that their research does not ask for verification in educational practice.

However, there has been a research tradition which has paid much closer attention to problems of educational practice than those traditions at which I have glanced so far. This tradition was historically known, particularly in Scotland, as 'experimental education', and it is represented and presented in condensed form in the United States by the areas covered in the First and Second Handbooks of Research on Teaching sponsored by our big sister, A.E.R.A.

Dominant in this research on teaching, which is largely concerned with curriculum, classroom process and teaching methods, is the so-called 'psycho-statistical paradigm'. (Fienberg 1977) This research paradigm originates in agriculture and its prophetic book is Ronald Fisher's Design of Experiments (1935). (I am personally indebted to my former colleague, David Hamilton, who pointed out to me how illuminating it could be to return to this source).

The paradigm is familiar to all of us here, I assume. The major breakthrough is the insight that random sampling is to be preferred to sampling judged to be representative because randomization allows the use of the mathematics of probability to estimate error, and to develop tests of significance. The structural threats to validity are comparability of the experimental and the control samples, known as the problem of internal validity, and comparability of both samples with a target population to which generalizations are to be made, known as the problem of external validity.

In agricultural research block and plot designs are used to randomize soil and aspect or other similar conditions, non-experimental variables such as, for example fertilizer, are controlled as is the experimental variable. (Incidentally, a Norfolk turkey farmer, working this kind of design in computerized experiments on factory farming, tells me that he

can get experimental effects in a block and plot shed of turkeys which turn out not to hold when replicated in an ordinary open shed).

In applying this research design to agricultural practice the objective is to make a preferential discrimination between one practical course of action and another. It is important to grasp that a similar design can be used to serve theory by providing a comparative discrimination between theories. The statistics may be seen as a decision-making technique which can be used to make decisions between theories or between practices. In agricultural research and in educational research the application to practice is more often direct than through theory. In education the design is used to attempt to discriminate between curricular specifications or between methods or between teaching styles, for example.

However, the yield of educational experiments which are cast in this mould is increasingly recognised as disappointing. I shall not trouble you with references. Rueful reflections abound. Most often experiments fail to yield statistically significant results. But when they do, there is still dispute on technical grounds. It is impossible to draw random samples in field settings in education; there is much interference from trait-treatment interactions; criteria of yield cannot be established as they can in agriculture.

Campbell and Stanley (1963) have looked at problems of sampling in particular relation to internal validity, Snow (1974) has subsequently reviewed them in relation to external validity. Both are cautiously optimistic. Walker and Schaffarzick (1974) are more pessimistic about criteria, at least in the area of curriculum, finding from a broad review of American work that curricula show up well when tested by instruments which favour them. Cronbach (1975), who has been hot on the heels of trait-treatment interactions, is now intensely sceptical, and even calls for humanistic approaches.

In this country the disillusionment has been less explicit, but the

reaction towards more descriptive, less statistical approaches is there represented as what has come to be called 'the illuminative approach'; Is the dominant paradigm in research in teaching proving ineffectual, I think it undoubtedly is. But why?

I want to argue that it is not because of technical shortcomings in sampling and statistical procedures that the classic designs are failing us. It is because of a misplaced conceptualization of the application of research to education.

First, let me assert the principle of individualization of treatment in education. I can illustrate this by comparing agriculture with careful gardening. In agriculture the equation of invested input against gross yield is all: it does not matter if individual plants fail to thrive or die so long as the cost of saving them is greater than the cost of losing them. If it costs  $x + y$  to hoe a field and the loss from not hoeing it is crop yield worth only  $x$ , then you don't hoe. This does not apply to the careful gardener whose labour is not costed, but a labour of love. He wants each of his plants to thrive, and he can treat each one individually. Indeed he can grow a hundred different plants in his garden and differentiate his treatment of each, pruning his roses, but not his sweet peas. Gardening rather than agriculture, is the analogy for education.

Now, assuming a perfectly executed experiment in teaching: the perfect sample, the perfect criterion measure, we should be able to conclude that if a teacher is compelled to adopt a single uniform procedure for all cases, he is better to adopt  $x$  than  $y$ . However, that condition is an important one, the discrimination of the relative effectiveness of two alternative procedures by the use of an experiment cast in the psycho-statistical paradigm is an effective guide to action only if a standard procedure must be used in all cases. At system level this implies uniformity of treatment in all schools irrespective of context. At classroom level this implies uniformity of treatment of all children.

Now there are many, I imagine, who, like myself, regard teaching as an art in which the teacher's skills are differentially applied as a result of diagnostic interpretation. Teaching is largely a response to the observation and monitoring of learning in cases. If this is so, then a crucial problem of the psycho-statistical paradigm as the design for a discriminant experiment is not simply that it deals in general prescriptions, but that it offers to guide teachers by overriding, rather than by strengthening their judgement.

In essence the design offers a measure of the probability that one procedure uniformly applied is superior by some criterion to an alternative procedure uniformly applied; it deals in terms of odds, and suggests where one might place one's bet. But it does little or nothing to explain differences of outcome between procedures.

Of this design I have said elsewhere: 'without understanding why one course of action is better than another, we could prove by statistical treatment that it is. The vision is an enticing one: it suggests that we may make wise judgements without understanding what we are doing and the difficulty of understanding what we are doing had been thought to be the barrier to wise action'. (Stenhouse 1978, p.28)

I could elaborate at a length quite inappropriate to this occasion a critique of evaluation of educational procedures by the use of the psycho-statistical model, and for today I must ask you to accept that such an elaboration would be reasonably coherent, if contentious. It involves the rejection of behavioural objectives as alternatives to hypothesis, the preference of norm-referenced over criterion-referenced tests (which are types of examination rather than research instruments), and the rejection of random sampling in research into areas of action. It argues that the breakdown of sampling necessitates a return to the study of cases. But it does not involve a rejection of quantitative approaches, of measurement or of statistical operations; some statistical techniques, such as time-series analysis, are obviously relevant to the study of cases observationally and experimental case

study can be equated with the N-1 experiment. Moreover, I believe sampling can be rethought. Of course, I do not claim to have done all this work. I am arguing for a shift of paradigm on the part of at least a proportion of researchers, for testing a new approach.

The key to this approach is the application of research to education through an appeal to teacher judgement. The assertion is that the improvement of teaching rests upon the development of the art of the teacher and not through the teacher's adoption of uniform procedures selected from competing alternatives.

If this is the way ahead, then we need a theory of teaching, and experimental results which do not make a theoretical contribution are of little worth. I am, of course, only echoing B.O. Smith (1961) and Ian Westbury (1971) among others; but I think I am extending the area of research in which the claim is being made, for I see tremendous theoretical potential in experimental action research based upon curriculum and teaching strategies, where their reference was mainly to observational research.

The point of view I am taking implied that research is best applied to education by producing theory which can enrich action. The action is the action of the teacher, and this implies that the theory of teaching must be understood by the teacher. Of course, this calls for greater research literacy among teachers, but it also calls for much more accessible research and theory. Since I believe that most educational theory is made more inaccessible to practitioners than it need be - not only because we researchers have a tendency to self-display, but also because we have personal intellectual needs which cannot entirely be sublimated into chess and crosswords - I think theory would actually be improved by being made more accessible.

Above all, however, a theory of educational action must be recognized as hypothetical. That is, its status as knowledge is

provisional and subject to revision. Moreover, its defects may be limitations of generalisation so that it serves as a framework within which theories of individual cases have to be developed. Such a theory cannot prescribe action, but can support only the development of experimental actions which test and refine or elaborate the theory. In short, an empirically grounded theory of educational action will require continuous revision and development through experiment, and much of that experiment will not be large-scale field experiment, but rather laboratory experiment.

There is thus a need for the development of educational laboratories. In them we shall have to control conditions so that we can simulate faithfully those of real classrooms.....

But wait a minute! That would seem to imply that wherever there is a real classroom there is a potential educational laboratory. Just so. The best designed educational laboratories are in charge of teachers, not of researchers.

The function of educational research in its application to practice is to provide a theory of educational practice testable by the experiments of teachers in classrooms. In a sense this calls for the development of the role of teacher as researcher, but only in a minimal sense. The basic desideratum is systematic inquiry; it is not necessary that this inquiry be made public unless it offers a contribution to a public theory of education.

Such a view of educational research demands of teachers the capacity to see educational action as hypothetical and experimental. Researchers on this view should disseminate to teachers a scepticism about research results and theories and hence a disposition to test them. Research should underwrite speculation and undermine assertion.

Research can be adequately applied to education only when it

develops theory which can be tested by teachers in classrooms. Research guides action by generating action research (or at least the adoption of action as a systematic mode of inquiry). Action research in education rests upon the designing of procedures in schools which meet both action criteria and research criterion, that is, experiments which can be justified both on the grounds of what they teach teachers and researchers and on the grounds of what they teach pupils. A systematic structure of such procedures I call a hypothetical curriculum. Such a curriculum is the appropriate experimental procedure through which research is applied by testing, refining, and generating theory in the laboratory of the classroom.

1 September, 1978.

REFERENCES:

- CAMPBELL, Donald T. & STANLEY, Julian C. (1963) Experimental and quasi-experimental designs for research on teaching. In GAGE, N.L. (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally 265-291.
- CRONBACH, Lee J. (1975) Beyond the two disciplines of scientific psychology. American Psychologist 30,2,116-127
- FIENBERG, Stephen E. (1977) Next steps in qualitative data collection. Anthropology and Education Quarterly 8,2,50-57.
- FISHER, Ronald A. (1935) The Design of Experiments Edinburgh: Oliver & Boyd.
- SMITH, B. Othanel (1961) A concept of teaching. In B.O. Smith & R.R. Emmis (Eds.) Language and Concepts in Education Chicago: Rand McNally.
- SNOW, Richard E. (1974) Representative and quasi-representative designs for research on teaching Review of Educational Research 44,3, 265-291
- STENHOUSE, Lawrence (1978) Case study and case records: Towards a contemporary history of education. British Educational Research Journal 4,2,21-39
- WALKER, Decker F. & SCHAFFARZICK, Jon Comparing curricula Review of Educational Research 44,1,83-111
- WESTBURY, Ian (1971) Problems and Prospects. In Ian Westbury & Arno A. Bellach (Eds.) Research into Classroom Processes. New York: Teachers College Press.