B E R A

* * *

PROCEEDINGS OF THE INAUGURAL MEETING

OF THE BRITISH EDUCATIONAL RESEARCH ASSOCIATION

* * *

April, 1974

University
of
Birmingham

FOREWORD

The planning committee of the Association asked me to compile the proceedings of our inaugural conference which was held at the University of Birmingham on April 4th and 5th 1974.

As you may recall the proceedings were lively and informative. There was a genuine concern to explore contrasting perspectives on research problems and these are reflected in the papers and reports presented here. Professor J. Nisbet's opening address sets out the main theme of the Conference and the papers by Gabriel Chanan and Professor M. Kogan examined the relationships between researchers and consumers. John Bynner's comments on research methodology acted as a springboard for vigorous discussion of conceptual and methodological problems in research. The symposia on classroom studies and evaluation both centred on paradigms of research and their underlying ideologies.

Finally, I should add that the papers were presented as informal contributions designed to generate discussion amongst members rather than establish definitive viewpoints.

G.A. BROWN
New University of Ulster

Contents   Page

EDUCATIONAL RESEARCH : THE STATE OF THE ART

Professor John Nisbet, University of Aberdeen

Address to the inaugural meeting of the
British Educational Research Association.

This attempt to review the present state of educational
research in this country is under three headings: growth,
trends and structure. I would have found it easier to deal
with the subject in a course of ten lectures - after all, the
Open University course E 341 on educational research requires
19 booklets, 11 radio programmes, 8 television programmes, 8
assignments and 1 project, just to provide an introduction to
the topic - but I shall do my best within the limits you have
kindly allowed me. The fact that I have a problem of compres-
sing the subject within one hour, is itself evidence of my first
point, the growth of educational research in recent years.
When I began teaching the topic in university, twenty-five years
ago, the text-books I inherited were Vernon's *The Measurement of
Abilities* (1940), Burt's *Mental and Scholastic Tests* (1921), Rusk's
Introduction to Experimental Education (1919) and Whipple's *Manual of
Mental and Physical Tests* (1912). Being very up-to-date, I intro-
duced Charlotte Fleming's *Research and Basic Curriculum* (1946), which
was the first example in this country of a compendium of research
which offered something more than just psychometrics. Nearly 20
years passed before we began to see the present flow of really
useful books summarising research findings in a general way,
like Thouless's *Map of Educational Research* and Butcher and Pent's
series *Educational Research in Britain*. In the meantime, we had to
rely on American books, journals and encyclopaedias of educational
research, and it is only in the last ten years that educational
research has established itself in this country as a topic in its
own right.

The setting up of this new association is another evidence
of the growth of educational research. Other writers - parti-
cularly, Wall, Taylor and, most recently, Vernon Ward - have
calculated how the national expenditure on educational research

has multiplied ten-fold since 1963. Whereas then only .01
per cent of all educational expenditure was devoted to
research, the figure is now about .1 per cent - or, for every
pound spent on education, one-tenth of a penny goes to develop-
ment, research and evaluation. This year, for the first time
ever, the number of SSRC postgraduate studentships in education
has reached three figures - 100 exactly. If that sounds a
small number, remember that seven years ago, there were 14
SSRC studentships in education; and even two years ago there
were only 65.

There has been a very substantial increase all round; but
one has only to see the present position in context to acknow-
ledge that there is still a long way to go. These figures
merely demonstrate that, until very recently, educational
research was a spare-time amateur activity for gentlemen of
leisure.

A more important form of growth has been the widening range
of types of research which can be reckoned as coming within the
scope of educational research. For many years in the past,
educational research was almost exclusively concerned with
educational psychology and testing; and though there were
eminent pioneers in the fields of the history of education,
the philosophy of education and comparative education, these
aspects of the subject were isolated and quite unsupported by
formal association of scholars. Recent years have seen a
remarkable growth of interest in these areas, and there are now
flourishing societies for each of the three aspects. The
emergence of curriculum development as an area of study and
research, has been the largest single element in the growth of
recent years, thanks to the work of the Schools Council. But
each of us has his own choice of factors which have led to the
multiplication of aspects of educational research: sociology,
educational technology, classroom observation, participant
observation, administration and management of schools, and so
on - a whole range of specialist disciplines.

The emergence of these specialisms, however, also carries
with it the danger that the study of education may split up
into less and less meaningful sub-divisions. There was a time
in Aberdeen, in 1960, when I taught everything in the Department,
alone - from Aristotle and comparative education to statistics
and testing. Needless to say, the teaching was not of a uniformly
high standard, and today in Aberdeen we have a staff of ten (and

even that is still a small department). But now we have to
have special meetings to ensure that there is some link between
the various aspects, and we require every member of staff to
do some tutorial work across all the boundaries. This is a
common problem in all disciplines, nowadays. In educational
research, in my view, it is particularly important that the
different aspects should not develop in isolation: the empirical
social scientist needs to draw on history, comparative studies
and philosophy. When Noel Entwistle and I wrote our report on
transfer from primary to secondary education, we had a section
on theoretical aspects, a historical section, a chapter on
comparative studies, an empirical follow-up study, and a small
piece of action research. Perhaps the correct solution is to
build up a research team which brings the different aspects
together. But I would argue strongly against the fissiparous
trend in current educational research; and if this Association
can bring together the groups which are now tending to go their
way, it will have served a useful function.

The second of my headings if 'Trends'. Among the various
sub-divisions I have been discussing, one of the most vigorous
in its growth is the sociological. For many years, educational
research was dominated by psychology: it aspired to scientific
precision in research design and hypothesis construction and
was preoccupied with measurment and statistical analysis. There

classical era to a modern one. But I don't see this trend
as dispensing with the need for rigour and precision. I don't
think so: there is a place for both styles, and certainly
students should be responsive to the merits of both, so that
they can recognise excellence and spot the flaws, whatever the
style.

Rather than see this as a confrontation, I suggest that
the trend of the past fifteen years has been the emergence of a
range of styles, which have added greatly to the power of
educational research methods. I used the word 'range', but
perhaps it is better described as a 'spectrum' -

Figure 1 : A Spectrum of Research Styles

1	2	3	4	5
Experimental method	Exploratory survey	Curriculum development	Action research	Open-ended inquiry
Empirical Educational science	Fact-finding as a basis for decision-making	new syllabus content and method. Field trials and evaluation	Interventionist	Grounded theory Participant observation Illuminative evaluation

	The agricultural model		The anthropological model	
Experiments to improve your products by manipulating treatments			Go and live there and see what it is like	

because a spectrum has no sharp boundaries, and also (if it is
not straining the metaphor) because you get white light by mixing
all the different frequencies!

One and Two represent the empirical tradition, which has a
strong Scottish-American flavour. The Scots who set up the first
educational research council in Europe in 1928 believed in it -
Thomson, Rusk, Drever, Boyd - and it can be traced back to Herbert
Spencer and to Alexander Bain who as a professor in Aberdeen
University was responsible for the teaching of Philosophy, Logic,
Rhetoric, English literature and language and Psychology, and also

wrote a book called Education as a Science in 1880. In fact, in
1946, in Aberdeen Training College (as it was then called),
the Department of Education and Psychology were combined into
a Department of Educational Science - an innovation which was
subsequently abandoned. The concept of a science of education
is based on the belief that educational problems should - and
can - be solved by objective empirical evidence, that precise
and accurate research can build up a structure of knowledge
which will generate new hypotheses and new experiments, until
the whole field is uncovered. This was a common aspiration in
the 1920's and 1930's, and it has an initial appeal to each
new generation of researchers. It is based on the faith that,
if only one could design a good enough experiment, with effective
controls, precise evaluative measurement and appropriate sensi-
tive statistical analysis, it should be possible to establish
objectively the one best method, the ideal curriculum, the
optimum period of instruction, the correct use of aids to
learning. Once these points have been established by experiment,
and adopted in the educational system, any change must be for
the worse. So the Scottish researchers tried to decide whether
it was better to teach children to add up a column or down a
column, whether in subtraction, the method of "equal addition"
was superior to the method of "decomposition"; whether "phonics"
was better than "look and say"; and today, the educational
technologists try to determine whether programmed learning is
better than traditional methods, whether television is better
than a live teacher, whether colour is better than black-and-
white; and in higher education, we are asked, "Which is better,
lecture or tutorial?" or "Is there any evidence to show what is
the optimum length of a lecture?" The educational science idea
keeps recurring: in 1964, Sanford, in *The American College,* wrote:

> "Practice in higher education, as in politics, remains
> largely untouched by the facts and principles of science.
> What our colleges do, tends either to be governed by
> tradition, or to be improvised in the face of diverse -
> and usually unanticipated - pressures." So Sanford
> envisages "A science of higher education . . . the notion
> that the field (of research) may ultimately be constituted
> as a body of fact and theory, a discipline of sorts, in
> which individuals become specialists".

But most educational researchers today no longer hold this faith in their power. Perhaps we have lost what was a guiding star to the pioneers, and a great source of strength, but it was always a myth. There is no one best method. Most empirical research studies are, to use Ashby's phrase, "miniscule analyses". This is hardly surprising: all science proceeds by "Miniscule analyses", and the building of a coherent theory is a very slow process on which we have only just begun. For the present, the problems which can be resolved empirically are relatively minor, compared with the major issues which require a judgment of values. Nevertheless, in these major issues, research has an important contribution to make, in defining objectives, in evaluation, in assembling relevant and adequate evidence on which to base our judgments.

And this takes me to the second category in my 'spectrum'. One of the distinctive developments of educational research in the past 20 years, has been the recognition that too often educational decisions are made without an adequate knowledge base. A major area of achievement is the fact-finding survey type of research, such as the national reading surveys every four years (approximately) since 1948, or the National Child Development study, or the Isle of Wight study, from which we can tell how many children are deaf, stutter, have behaviour problems, are left-handed, wet the bed (and at what ages) and so on. Most of the major educational reports - Crowther, Newson, Robbins, Plowden - have been accompanied by extensive surveys, and are much the better for it.

But there is also a potential weakness in this kind of research, for much of it is news rather than science:

> Children in the North of England watch television more than children in the South of England;
>
> 18% of junior school teachers do not know how to begin teaching reading;
>
> 30% of medical students live at home.

In the absence of theory or hypotheses, there are useless pieces of information. To quote on reviewer (Holmes, BJEP, 1972)

> "At first sight, . . . research might appear to be thriving.
> But this impression results from the use of the term

'research' to cover work which might better be designated as . . . development, survey or information. This umbrella usage tends to obscure the fact that controlled evaluative research . . . is rare."

My third category is curriculum development, which is the largest single growth area since 1960. It occupies a middle position between my two extremes, drawing on theory and survey, using experimental work and field trials, prepared to venture into open-ended inquiries - as is appropriate for practical-oriented development work which attempts to bridge the gap between theory and practice. If there is a danger of weakness here, it is - (and I hope you are beginning to see the underlying theme of this address, the moral of my tale) - that curriculum development is inclined to isolate itself as a special new kind of discipline, with its own specialists, its own techniques, its own jargon and even its own funding organisation, instead of recognising how much it has to gain from well designed experiment and evaluation on the one hand, and interventionist and exploratory-type studies on the other hand.

An so we come to the last two categories, which offer a promise of transforming the whole field of educational research, if only we can prevent them from being used by people as a short-cut to 'instant research', to avoid the trouble of thought and planning, or to cover up a lack of knowledge or a willingness to submit one's ideas to the test of hard evidence. It was Michael Young who pointed out the calamity of the gap between research and innovation: there is innovation without research - new ideas based on hunches, never tested objectively; and there is research without innovation - academic studies which make no impact and are unintelligible except to other researchers. Bringing research and innovation together in 'interventionist-type studies' gives 'action research', in which research monitors change, research is a guide to action, and the results of action are a guide to research.

Halsey's review of action research in Chapter 13 of *Educational Priority*, Volume 1, is as lucid an analysis as is to be found; and I doubt if there is much I can add to what is said there. As I see it, the problem is to keep research alive beside its dominant and vigorous partner; but the action men in the EPA studies saw

the problem as winning freedom for action from the cold restrictions imposed by the researchers. The tension exists between the two concepts, action and research: action has all the pupular qualities - commitment, involvement, belief, enthusiasm; the qualities needed in research have a more limited appeal - detachment, suspension of belief, scepticism. Or, to use the vocabulary of Elisabeth Richardson: for action, there must be loyalty, and loyalty is 'a collusion to maintain the pretence of infallibility'; but research requires a tolerance of heresy, 'a willingness to submit the most sacred ideas to the test of reality'.

Action research developed as a protest against the scientific detachment of traditional psychological and psychometric studies. This protest has now been carried to its logical extreme in the last of my five categories of research. In the anthropological model, to understand the educational process, to do any effective research in education, one must see it from the viewpoint of the learner. Traditional empirical research, especially when it involves testing, experiment and statistical analysis, start from our assumptions, our framework of thought, and it imposes that framework on what we innocently call the "subjects" in our experiments. Not surprisingly, the framework usually does not fit; and so we adjust our control mechanisms until we have a situation where we can use our preconceived models - and, not surprisingly, this kind of research produces results of limited value and limited application.

So we have a new style of research, and it is one which we must come to terms with. Just as psychology dominated the 'educational science' style, so sociology dominates this opposite extreme. Here it is important to build constructs on the basis of open-ended inquiry. The case study reveals the unique features of a situation. Participant observation enables the observer to get inside the skin of a situation, instead of studying it in a detached way. Grounded theory is built up from observation, not imposed a priori. The descriptions used by different participants to explain their experience provide an exploratory tool; and thus theory is grounded in the everyday life of the people who are being studies. Illuminative evaluation uncovers the nature of what has happened: it does not prove that x is better or worse than y - (how could it be? they are different, and comparison is irrelevant).

This style of research, like the others, has its potential weaknesses. It can be an excuse for indiscriminate data collection, for tiresome transcripts of trite interview exchanges - research without ideas, which is research without interest, the anecdotal model, rather than the anthropological model. But as its best, this is a highly sophisticated and perceptive style of research (and I wish we knew how to teach the skills on which it depends). It can be even more demanding than the relatively straightforward laboratory experiment, and it may require the use of complex statistical analysis or elaborate procedures like the repertory grid.

We are often presented with this style of research as a challenge to the traditional model. To some extent it is; but my position is that no one of these styles is "right", and none is altogether "wrong". The most effective research employs a variety of strategies, across the spectrum.

I have splken about growth and trends: my final heading is structure, and this includes both the organisational structure of research funding and the infrastructure of research support.

There is clearly a need for a meeting place like the British Educational Research Association, to bring interested parties into effective communication with each other. But many people would want to go further. For example, the Universities Council on the Education of Teachers, in a research policy document, stated:

> "There should be consideration at national level of the
> possibility of establishing better machinery than at
> present for the identification and discussion of research
> priorities and the co-ordination of research policies
> and initiatives".

Can we envisage the creation of an organised and integrated structure for educational research, for planning, funding, monitoring, for developing research support services, overseeing the provision of research training and financing postgraduate students? I confess that the creation of such a structure is a grand ambition - or, perhaps a pipe-dream - which I have had myself from time to time. I suspect it was in the minds of some when the Schools Council was set up in 1964, and again in 1965, when the Educational Research Board was established. In both cases, if people had such expectations, they have been disappointed. But from time to time there is talk of creating an

Educational Research Council, a supreme body, which would oversee the work of the wide range of institutions concerned with research and development in education, integrating their activities and determining their priorities.

None of the bodies involved in research funding is enthusiastic about this kind of monolithic structure. Perhaps they are merely defending their vested interests, but the argument against it is well made in W. C. Radford's recent report, *Research into Education in Australia,* 1972.

> "Co-ordination is useful, provided it does not throttle intellectual independence and initiative. In the complexities of the social sciences, complete co-ordination of research would require omniscience and should never be attempted." (footnote, p. 6, from Conrad)

> "The development of a subject is to a large extent a gradual uneven growth . . . and the most that can be done is to ensure that the system . . . discourages the growth from being too uneven or too gradual." (ditto, from Cunningham)

> "Let me make clear immediately that I do not believe in the laying down of priorities in research by a central body . . . . My reason is simple. Such a laying down of priorities to me implies an impossible omniscience, and lays up trouble for itself . . . . Provided that those engaged in research develop adequate channels of communication between themselves . . . I believe there will not be any greater gap between the nature of problems and the information from research available to solve them, than there would be were there to be a central determination of a limited number of priority areas in which alone study would be supported. It is as well to remember that, not very many years ago, 'education as investment' and 'manpower studies' loomed very much larger as matters of research than now seems warranted by later experience . . . . Had the major part of the research apparatus swung over to such studies in 1965 or 1966, a good deal of work now known to be more valuable would not have been done."

I am reminded of Berlyn's story of the response of an imaginary advisory council in 1810, asked to forecast the development of the transport system. "One thing", they concluded, "has stood the test of time over several thousand years: the horse has come to stay. Authorities as diverse as Genghis Khan, Dick Turpin, Julius Caesar and Buffalo Bill, all agree on one thing, from ong experience, that there is no better way of getting from one place to another than on a horse."

So the attitude of the Educational Research Board has been mainly responsive. 'Responsive' does not mean waiting for others to make suggestions: it means being ready to respond to imaginative ideas, and resisting the temptation to impose one's own ideas. Few other bodies are prepared to do this: few are able to do it with public money. Mrs. Thatcher in 1970, expressed the direct opposite view for DES-sponsored research:

> "There was clearly only one direction that the Department's research policy could sensibly take. It had to move from a basis of patronage - the rather passive support of ideas which were essentially other people's, related to problems which were often of other people's choosing - to a basis of commission. This meant the active initiation of work by the Department on problems of its own choosing, within a procedure and timetable which were relevant to its needs. Above all, it meant focusing much more on issues which offered a real possibility of yielding useable conclusions.

This is an appropriate view for a government department, but it is also appropriate for a body like ERB to be prepared to operate outside the limits of established policy. The 1973 report of ERB said:

> "When the Board reviewed the problem early in 1973, its decision was to reaffirm the 1971 policy statement, that whereas 'the other major bodies are chiefly concerned with policy-oriented and basic research is open to question, and there is no implication in the statement quoted that SSRC is interested only in 'pure' social science inquiries the intention is to bring out an emphasis on the theoretic contribution which each approved project should make as at

least part of its results. This is an aspect of educational
research which does not seem to be adequately dealt with by the
other funding agencies."

There is a place for both approaches and a need for partnership between
them. There is a place also for each, occasionally, to do something of
the other's role: thus, the ERB has embarked this year on the task of
instituting a programme of research in pre-school education, and this
will be developed in conjunction with the national programme of DES,
DHSS and the Scottish Education Department.

There is, however, one aspect of the organisation of educational
research where there is an obligation - an urgent need - to undertake
positive initiatives. This is to build up an extensive infrastructure
for research and development. William Taylor has argued this point
persuasively in writing and in speeches: the most effective way to improve
the quality of educational research, he says, is to build an adequate
"research floor" - funds, equipment, personnel, procedures, training
programmes, communication, information retrieval, and so on. To take
only one example, information retrieval, the Educational Research Board
has recently brought together the representatives of fifteen different
organisations to agree on a common format for abstracts of research; and
with feasibility studies begun in the NFER and the Scottish Council for
Research in Education, working in collaboration with the Council of Europe
EUDISED Thesaurus of educational terms, we have the basis for a possible
information retrieval service covering the whole European scene, to match
the ERIC system on the other side of the Atlantic.

So, under my last heading of structure, there has been some progress
towards a fairly open system, and towards more effective supporting
infrastructure. If the new BERA is also a fairly open organisation, and
avoids the danger of domination by any one group that believes that it
alone possesses the key to understanding, then there is undoubtedly an
important part for it to play!

To summarise, I think that recent years have seen a move away from
the naive idea that problems are solved by educational research: that is the
old "educational science" idea, and it is a myth. Educational research can
strengthen the information base of decision making; its procedures of enquiry
and evaluation - inject rigour into the flabby educational thinking that has
satisfied us for too long in the past. The most important contribution of

research is, I suggest, indirect. This is important: in one sense, educational researchers are unacknowledged legislators of the next generation. As Taylor says in his recent book, Research Prospectives in Education, a primary function of research in education is to sensitise - to make people aware of problems. Also, in assessing the achievement of educational research, we have to consider its effects on the attitudes of those who teach. Vigorous research activity or, to use a less pretensious title, investigation into teaching and learning, sharpens thinking, directs attention to important issues, clarifies problems, encourages debate and the exchange of views, and thus deepens understanding, prevents ossification of thinking, permits flexibility and adaptation to changing demands. Research of this kind aims to increase the problem solving capacity of the educational system, rather than to provide final answers to questions or objective evidence to settle controversies. On this view, educational research is a mode of thinking rather than a short cut to answers. In the long run, the real influence of educational research is through its effect on the attitudes of those who teach.

THE POLITICS AND FINANCING OF RESEARCH

Notes of a Talk Given by Professor Maurice Kogan, School of Social Sciences, Brunel University*

"Politics" has several meanings including the theory of the nature of the state, the relationships of individuals and institutions with it and is also concerned, most fashionably, with a form of voyeurism which is at best micro.history. This Introduction will be concerned more with the structural relationships which educational research creates - the government of education research, than the politics.

The following questions might be posed.

First, what are the ascribable functions of educational research. Second, who might get what out of them? Third, who does, or should have, the authority to fund, or to veto funding, to allow or to refuse access, to educational research? What might be, then, the main roles involved in a rational research structure, and what should be the nature of the main relationships between them? And - politics at last - how might the resulting role structure relate to the less systematic networks of power and influence and sentience in the educational service?

Functions

Let me remind you briefly of the classic differentiations between "conclusion-oriented" and "decision oriented" inquiry to be found in Cronbach and Suppes' "Research for Tomorrow's Schools". Or another, the Rothschild Report's differentiation between the customer (government department) and the contractor (the research councils or the researchers) which emphasises heavily what Cronbach and Suppes would call decision or mission-oriented research. There are a second set of distinctions which, on the face of them, do not relate so much to criteria of usefulness as to their epestemological status - the sort of continuum laid down by the U.S. Office of Education and discussed in the 1971 O.E.C.D. Report, as between fundamental or basic research, applied research, development, dissemination and installation. These two sets of categories overlap.

---

* The speaker wishes to make it plain that this talk was prepared on the understanding that it was an introduction to an informal discussion and was not a researched paper for publication.

Authority

Yet distinctions are necessary if we are to ascribe
authority to make decisions about who creates research. All
the same, the distinctions are difficult to sustain. Differ-
ences between fundamental and applied research are hopelessly
complex in an area such as education where there can be no
study which either does not derive from the field or usefully
relate to it. One commentator (Eide) has observed that
fundamental and applied research differ not in the type of
problems being researched but in the autonomy that the researc
actually or potentially should have.

Types of Research

What are the functions - how can one describe the role of
research? Some writers such as Glennister and Hoyle have
usefully provided a shopping list. I am more concerned with
stating a role description for educational research and develo
ment and for ascribing consumers to each type of research and
development before discussing their government:-

First, research could provide a knowledge base against
which policies, the authoritative allocations of values could
be tested. This would not determine policy but would help
policty makers bring into a disciplined framework what is
knownabout, for example, the impact of nursery education upon
different pre-school age groups and different social classes
within different environments, or what research might tell us
about the effects of raising the school leaving age. Some of
us think this might be useful to the D.E.S. and to local
authorities. Such a policy backing up service by researchers
would help identify the undiscovered areas.

May I emphasise, however, that such an analysis may not
make out a claim for research. It may, on the contrary, show
both administrators and researchers that some of the more
important issues are not usefully researched, but are suscepti
ble to disciplined administrative analysis rather than academi
disciplined enquiry. Or that, of course, an issue is best
treated as requiring a political rather than a more systematic
reaction. The links, however, between administration as in
programme analysis review and research need to be identified
and strengthened.

Secondly, research is concerned with evaluating professional educational practices, and developing better ones. Curriculum development, relation of assessment schemes to learning, the timing and length of the educational processes and their sequences, sizes of groups in education. There is a whole area of research at least as much of interest to teachers, their educatiors and local authorities as to the distributors of resources.

Thirdly, there are researchers which concern the function of education in its widest social context, as a distributor of life chances, as a user and creator of manpower, as a social control mechanism, and a system which relates to other social welfare provisions as a provider of increased freedom and auto-nomy for individuals. These researches link with the first and the second kind but bring the researcher into a different authority and power structure. For here he might be required not so much to service the educational system as to criticise it, test its objectives, help it change them, and, indeed, strenghen those who wish to have it changed, perhaps in opposi-tion to the professional educators.

Fourthly, there are studies of educational organisation which are becoming far more prominent than they used to be and range from, say, the Bristol study of inspectorates or the internal organisation of a secondary school, to studies made, not in this country, for O.E.C.D. but in the U.S.A. or in Turkey about the structure of the whole educational system. These studies might or might not be of use to and requested by central or local government. Equally they might be undertaken by those who regard it as important that there should be licenced criticism and testing of governmental systems. Here the question of relationships, of politics, becomes critical.

Fifthly, the ongoing process of organisational development. This calls for a different type of researcher, an ongoing collaborative relationship, unless it is simply to be a system buster.

This is a somewhat random list of researches but it helps me to emphasise that the question applied to Blau and Scott to types of organisation, "Cui bono?" is relevant here. None of

us need doubt that there is plenty of work for researchers
to do and that the system needs it. But we cannot impute the
right status, or the right mode of governing and working
relationships, to the research unless we are clear as to the
imputed beneficiaries of it. And here, we are in far more
difficulty than our colleagues researching on, say, health
or housing services. The health service are not dealing with
so varied and potentially conflictual set of perceptions of
what the whole process is about. Everybody believes in mendi
broken arms.

Authority and Resources

My third question was where the authority to fund or
sanction research rest, and where should it rest? The formal
answers are easy to find. The D.E.S., the S.S.R.C., the
Schools Council, individual local authorities and a few
private foundations are virtually the only sources of funds
in this country. There is no doubt that central government
is and will remain the main funder of education research.
Much as we may rage against the coming of the night, this is
the fact. As demands for university places shrink, and as
the love affair between the universities and Whitehall become
increasingly jaded, university teachers ought to regard their
own time as increasingly available for research. I have hear
Roy Parker protest strongly against the assumption that all
research has to be done by a large unit led for a half a day
a week by a professor and with a team of ten research assis-
tants all jostling for tenure and a place in the sun with
them. We have all found it necessary to create this sort of
pattern, but it may well be that the lone senior academic
ought to be restored to his place so that concept forming car
begin to push data collection to one side, particularly as we
come forward to tackle problems with which the education ser\
is rife. Wright Mills' playfulness and just general creativi
and intellectual fearlessness are among the things that we nc
seem to be lacking. Our only way out, if we wanted one, woul
be to use individual free time within university departments,
or to look towards the private foundations. But they, and
particularly the Nuffield Foundation which has been one of tl

most imaginative in the world in looking for points of growth
in the education service, simply do not have the resources
to support more than a tiny proportion of what must be found.

Central government is using its funds through two princi-
pal routes. The first is through the D.E.S. and Scottish
Office research funds. The second is through the research
councils and, particularly in our case, the S.S.R.C. and the
Schools Council. The distinctions between them are well
known and can be briefly typified. The D.E.S. is enjoined
to seek a strong customer-contractor relationship with those
whom it funds. It is not empowered simply to act as patron
to the development of knowledge for its own sake but have to
have a policy useful pay off for what it funds. Moreover,
much of the research council budget is to be devoted to satis-
fying initiatives from the customer departments. And, further-
more, one suspects that pressure is on the research councils
to apply the criterion of policy usefulness to much else in
their research programmes as well.

But given the fact that the money comes from central
government and that the days are now passed when they have no
clear interest in the purposes on which it will be spent, we
still have other important characteristics of any public
research programme which both mitigate and complicate the
relationships between government and the research community.
In the first place, government and the research councils are
advised by members of the research community. Indeed, so
much so that administrators in the civil service are some-
times not too strong but too weak in facing up to research
interests. They start, very often, with a lack of conviction
about the role of research in any of the functions which I
began my talk by stating. But once they have become convinced,
are they not likely to simply capitulate on the assumption
that the social scientists' contribution, once acceptable in
principle, is likely also to be right? I have before said
that we need courageous and more entrepreneurial bureaucrats
in charge of the welfare state and I would like to see them
roubust enough to state what are the main policy areas about
which they are uncertain, and to enter into dialogues with
the social scientists to see whether they can provide any of

the answers, or change the framework in which the answers might be made. We do not suffer from a bullying administrative and ministerial attitude towards research but a somewhat withdrawn and unformulated system against which to play our role. The link with the policy analysis is not made. In fact, everything now points to social scientists coming out of their corners and asking for a far more overt and frank an explicit system by which government is going to make use of social science so that it can be challenged, improved and cau to run fruitfully.

Here, might I say, the one department which has done thi in any sense at all, is the Department of Health. Their chie scientists' organisation, with its perhaps elaborate set of advisory committees and panels, and its overt relationships with the policy divisions and research management divisions - all published in the Nuffield Provincial Trust's recent versi of "Portfolio for Health" can be criticised on several scores But the Department created the system, promulgated it for wid discussion, and now is openly implementing it. It will suffe no doubt, from all of the penalties of having somewhat presti gious medical and social scientists placed in the position of judging their peers but this is unavoidable, and at least we know who is doing the judging.

I now turn to the last of my questions. What, then, would be the best role structure, and what would it be up against?

In the first place, our interest goes further back than the research and development system. For we are concerned wit disciplined enquiry into the educational service. We are concerned with ways in which the educational process might fruitfully be improved, and ways in which their impacts on the whole society might be enhanced. But these words are meaningless unless they are tied to explicit statements of values, objectives and programmes. So the starting point of any research and development system which is to be both demo cratic and effective is, in fact, a statement of educational policy. Hence the importance of linking the study of the knowledge base to the development of programmes, and hence

the argument for making the D.E.S.' programme analysis review part of an open system. Social scientists could help with this, not necessarily on D.E.S. grant, but as academics able to apply their minds to what is proposed, if it were made more overt. There are various panels being developed by the D.E.S., for example, the development of nursery education. But do we know who these people are and what expertise they are supposed to bring to the process? In other words, I think we need a chief scientists' organisation parallel to that established by the Department of Health. This should enable the D.E.S. to catch the best opinions about where research and development can best be applied, and to what purposes.

Next we have the terribly sensitive issues of how far educational research is to be tolerated - whenever it gets near important structural questions such as the authority of administrators over teachers, or the role of inspectors, or ways of evaluating education performance, with the teaching profession and the local authorities. In the past, too, the N.U.T. have been far too sensitive about potential researches. Obviously, the more trivial and badly regulated enquiries are simply a nuisance and a work creator for the schools but responsible researchers need to be given access, and the schools and local authorities should try to be welcoming rather than defensive about them. For there is here a political phenomenom with which researchers have to contend. The education service is more strongly consensual than any other part of the public sector. It has a liberal consensus which is reinforced by the training systems, by local authorities, by H.M.I.s and by successive ministers. In Edward Boyle's words,

> "Overwhelmingly the biggest number (of fresh
> policies) originated from what one broadly
> ccalls the education world ..... from the logic
> of the education service as it was developing".

But the way in which the schools are run, who makes decisions about them, the type of society that the teaching profession is moulding through its educational function, are matters that affect the whole society. This being so, it is really the role of the social scientist to bring up for public

testing the way that schools are managed and the purposes foɪ
which they are funded and staffed by the public.

All of this points, then, to a stronger and more explic:
structure to be created by the D.E.S. in educational researcɦ
It should freely draw on advice and implicit sanctioning froɾ
the education service at large, the local authorities and thє
teachers, as well as social scientists, but the assumption sɦ
be that within braodly agreed frameworks of research, custome
oriented research will not be tramelled by defensiveness. Aɾ
at the same time, if government wants social science to be
worth anything at all it will, at the same time, do something
to ensure that there is at least some free money available fc
researches and synoptic assumptions about educational policy.
Some of the American government contracts allow for "free
monies" so that their contractors will be able to do their ow
thing at the same time as they are helping the government. I
this country, it may well be that contracts should be writte
in that way. But I suspect that the right way through is to
create a distinctive margin within U.G.C. funding to the depa
ments for research. That margin could be adjusted quinquenni
if it becomes obvious that the department concerned is not ma
a lead in its own area.

Lastly, the researcher in all this. In all of these mat
we have to take note of his role. It is ambiguous. For the
most part he is expected to be a service giver, a giver of
information and advice that can be ignored. If he attempts
to become more than a service giver and be a decision maker o
even work close to decision making as in say, local authority
R and D work, his right to provide informative criticism is
diminished. The politics and government of educational resea
need, then, to be constructed so that he is able to relate
usefully to those who fund him whilst not losing the ability
to test the system that funds him. Where this leads us as a
group is not clear. But ambiguity is in any case deep within
the role of the teacher as well as the research. We all have
to be helpers, critics, evaluaters, testers and legitimisers.
I can see no way out of this.

POLITICS AND FINANCING OF EDUCATIONAL RESEARCH

Speaker : Professor M. Kogan

Professor Kogan began by giving an analysis of some of the problems in this area. This structure will be used to summarize some of the points made in the two-hour discussion period which followed.

Potential Functions of Educational Research

Five potential functions were identified:-

i)   to provide a knowledge base for policy decisions;
ii) to evaluate educational practice;
iii) to monitor the wider social control aspects of the educational system;
iv) to study particular organisations and sub-systems;
v)   action research in relation to development.

For each of these potential functions it was anticipated that a different network of relationships between researchers and policy-makers or practitioners would develop.

Discussion

i)   Who defines the research problems?

With official funding bodies, there may be some reluctance to provide funds or access to researchers when the findings may be critical of policy-makers. Differences amongst various government departments were noted in this aspect and it was suggested that these differences might be associated with the political implications of the research problems.

The point was made that research was not synonymous with funded research. The need for some researchers to be independent and to define their own problems was stressed. The extent to which the traditional research note of individuals within Universities had declined and required re-examination was raised. It was also suggested that researchers were somewhat timid and lacking in spirit in formulating their proposals. The influence

of policy trends, the perceived bias of fund-granting networks,
and the prevailing political and social climate of the times
on researchers' awareness of what projects might be likely to
be funded was indicated.

The difficulty of broadening the basis of the decision-
making of funding bodies was mentioned.

ii)   Who interprets the findings?

The way in which the same information is presented, accor
ding to the purpose of the group which presents it, was
illustrated using the limitations of the 11+ selection and the
introduction of comprehensive education.

Role of Central Government on Funding Body

Professor Kogan suggested a number of issues that require
to be examined:-

i)   need for central government to be clear about the kind o:
      research it should be funding: even in areas where the
      policy implications were clear it may be that empirical
      research is less necessary than other forms of analysis;

ii)   need for a new class (or approach on the part) of civil
      servant - "the entrepreneurial bureaucratic" - to relate
      the researchers and to confront them when necessary;

iii) need to examine the organisational structures concerned
      with research and planning. The need to open up the
      system was stressed.

Discussion

The political nature of educational policy decisions, tl
complexity of the issues about which decisions are necessary
and the limitations of research conclusions were felt to be
important considerations. While opening up the systems and
making explicit the process whereby decisions were made may
advantageous to researchers, the costs to and difficulties o
the central government were appreciated. It might be unfort
if openness led to attempts being made to systematize what w

difficult to systematize or to elaborate public relations exercises. Evidence about the outcome of opening up complex processes was lacking.

Relationships between researchers and those likely to be affected by the findings

Ways in which policy-makers could relate to researchers were considered, given the sensitivity of decision-makers to usual findings. The education service is, perhaps, too consensual and while it was necessary for it to be able to defend itself in the face of negative forces, Professor Kogan felt that ways in which it could be exposed to research findings should be examined.

Discussion

The restrictive nature of policy-making institutions was not accepted by all participants. The examples of Sweden and Russia were suggested as worthy of examination, but it was emphasized that all relationships between policy-makers and researchers could not be charted. The nature of the communication that might be established was also important and the extent to which mutual awareness of the difficulties and problems of both groups could be developed.

PROBLEMS OF METHODOLOGY IN EDUCATIONAL RESEARCH

Introductory Speaker : John Bynner (Open University) who provided three subheadings for his talk -

   (a) philosophical bases of research,

   (b) research designs, and

   (c) data analysis.

Summary of points raised in discussion - 5th April 1974.

1. Bases for Research

  1.1 A variety of purposes were envisaged for educational research:

  1.11 To investigate educational claims and the arguments advanced to support them.

1.12   To say what ought to be done.

1.13   To say what is possible.

1.14   To get people to change.

1.15   To generate 'useful knowledge'. To produce conclusions 'leading to effectiveness and efficiency' (It was pointed out that this implies a knowledge of what ought to be done).

1.2 A number of research strategies were suggested:

1.21   A 'scientific' strategy, based on a 'hypothetico-deductive' method, which should 'embrace other approaches within it'. (The strategy preferred by the introductory speaker.)

1.22   A 'phenomenological' strategy which involves a 'description of the contents of consciousness without the aid of a theory of the external world'. This description precedes attempts at empirical research and involves a redefinition of what it means to be 'scientific'.

1.23   A strategy which allows the scientific and the phenomenological to co-exist. Educational questions refer to private as well as to public universes of discourse, e.g. it is necessary to find out by what criteria teachers make judgments.

1.24   A strategy based mainly on survey methods, on the grounds that educational research is in a 'pre-theoretical' phase, requiring description rather than rigorous hypothesis testing.

2.   Practical and Ethical Problems

2.1 For whom is research done, and what do those who are the subjects of research get out of it? How are the problems over access to samples for research to be overcome?

2.2 What are the implications of policy decisions about the kinds of research which should be funded?

2.3 To what extent is it justifiable to set up experiments which affect the education of individuals, e.g. the implementation of experimental curricula.

3. <u>Methodology</u>

3.1 What are the benefits and drawbacks of random sampling? Is it true that it 'always pays dividends', or may there be situations in which behaviour is so context-dependent that randomisation produces meaningless results?

3.2 What problems result from the widespread availability of computer 'packages' (e.g. SPSS), which allows complex analyses to be readily engaged in? Should efforts be made to confine their use to those who 'know the meaning of data?' How could this be done when the 'meaning of data' is, in any case, something dependent upon which a number of competing ideologies of research is supported?

3.3 There was widespread support for a 'multi-disciplinary' approach to educational problems. Contributions can be made by anthropology and socio-linguistics.

3.4 Popular concepts (e.g. 'motivation') could be distorted in the interests of a particular research methodology. Such notions might be better investigated from the teacher's point of view. The elicitation of 'accounts' might lead to generalisation.

3.5 Much is to be gained by planning for enduring relationships between teachers and researchers, involving feedback and collaboration in the definition and analysis of problems.

Reporter: W.A. Reid (Birmingham)
Chairman: R.C. Whitfield (Cambridge)

B.E.R.A. INAUGURAL MEETING

Revisions to the published draft constitution

1.   Article 4, second sentence, replace by the following:

Except for the 1974 A.G.M., four positions on the Executive
Council will be open for annual election comprising the Vice-
President and the three members without portfolio. The offic
of President will be filled annually by the outgoing Vice-
President. The following five officers will be elected for a
three-year period:

> The General Secretary
> The Treasurer
> The Membership Secretary
> The Conference Secretary
> The Publications Editor

2.   Article 7 to read:

Proposals for changes to the Constitution shall be submitted
to the Executive Council who shall communicate such proposals
with their recommendations to the voting membership by post.
Voting members unable to attend the A.G.M. may cast their vot
in writing by proxy through the General Secretary not less th
three days prior to the A.G.M. Constitutional changes requir
the support of 65% of the votes cast before adoption.

R.e.W./5.4.74

PROBLEMS OF METHODOLOGY IN EDUCATIONAL RESEARCH

John Bynner, The Open University

I do not propose to give a paper on this topic but rather point to three main areas which may stimulate and guide discussion. These may be roughly described as the philosophy underlying research, research design and data analysis.

Research philosophy

Perhaps the most important point to make in this field is that educational research is concerned with the generation of useful knowledge. This is its rationale but of course the term 'useful' raises questions such as 'in what way' and to 'whom'. There are also questions about what counts as 'good' research, there is conflict between those who favour hypothetico-deductive strategies of enquiry for developing theory which emphasises the disproof of hunches by recourse to observation and those who prefer phenomenological approaches which emphasise subjective meanings of events. At this stage I will not attempt to argue the merits and weaknesses of these disparate views but I do suggest that we need to reconcile these viewpoints within one framework.

Research Design

This leads me into the area of research design. For me, the value of phenomenological approaches lies in their capacity to alert attention to important variables and to conceptualise variables. They are particularly useful at the pre-theoretical stage of describing and classifying phenomena with a view to generating hypotheses rather than testing them and they are also useful at the exploratory and pilot stages of survey research when researchers and respondents should be working together to develop research hypotheses.

Beyond the pilot stage, phenomenological approaches are no longer as useful as the correlational and inferential procedures of psychometrics which, if used sensitively, can rigorously test operationalised hypotheses and explore general relationships. Here I would make a plea for greater use of modern developments

in psychometric theory such as Cronbach's contributions to con-
cepts of reliability and Campbell and Fiske's concepts of
convergent and discriminant validity and a greater knowledge o:
the technology of sampling - particularly random sampling pro-
cedures.

Data Analysis

I suppose the central aims of data analysis are to accurat
summarise a wealth of information and to produce elegant parsi-
monious explanations. This inevitably involves elucidating
complex relationships between variables and this is usually
carried out on high speed computers. Easy access to the power-
ful methods has disadvantages as well as obvious advantages.
For example, techniques of factor analyses are often used inap-
propriately. Access to powerful methods with little recognitio
of their strengths and limitations may generate substantial yie
of useless knowledge.

Just as there is a need for more sophisticated methods of
experimental design so too there is a need for improving and
extending researchers' knowledge of survey analysis techniques
developed by Lazarsfell, Davies and Gochran and of econometric
techniques such as path analysis. There are advantages in usin
Boyesian methods of inference. In addition we need to look
again at the advantages and disadvantages of parametric and non
parametric methods.

Conclusion

The problem areas touched on in this paper are common to al
social sciences but perhaps in Education more than any other
field there is a need to combine the experimental rigour and
measurement sophistication of psychology with the strategies of
hypothesis formulation and survey technology of sociology and
the subtle elusive but often illuminating approach of the
phenomenologist.

1. **Bases for Research**

    1.1   A variety of purposes were envisaged for educational research:

        1.11   To investigate educational claims and the arguments advanced to support them.

        1.12   To say what ought to be done.

        1.13   To say what is possible.

        1.14   To get people to change.

        1.15   To generate 'useful knowledge'. To produce conclusions 'tending to effectiveness and efficiency' (It was pointed out that this implies a knowledge of what ought to be done).

    1.2   A number of research strategies were suggested:

        1.21   A 'scientific' strategy, based on a 'hypothetico-deductive' method, which should 'embrace other approaches within it'. (The strategy preferred by the introductory speaker.)

        1.22   A 'phenomenological' strategy which involves a 'description of the contents of consciousness without the aid of a theory of the external world'. This description precedes attempts at empirical research and involves a redefinition of what it means to be 'scientific'.

        1.23   A strategy which allows the scientific and the phenomenological to co-exist. Educational questions refer to private as well as to public universes of discourse, e.g. it is necessary to find out by what criteria teachers make judgments.

        1.24   A strategy based mainly on survey methods, on the grounds that educational research is in a 'pre-theoretical' phase, requiring description rather than rigorous hypothesis testing.

2. **Practical and Ethical Problems**

    2.1   For whom is research done, and what do those who are the subjects of research get out of it? How are the

problems over access to samples for research to be overcome?

2.2   What are the implications of policy decisions about the kinds of research which should be funded?

2.3   To what extent is it justifiable to set up experiments which affect the education of individuals, e.g. the implementation of experimental curricula.

3.   Methodology

3.1   What are the benefits and drawbacks of random sampling? Is it true that it 'always pays dividends', or may there be situations in which behaviour is so context-dependent that randomisation produces meaningless results?

EDUCATIONAL RESEARCH AND ITS AUDIENCE
Gabriel Chanan (NFER)

The free-market situation

If we start from the premise that all educational research results are entitled to effective dissemination, it seems to me that we have little ground for complaint in the present 'free-market' situation: individual reports, journals, projects are competing for attention, competing with each other and with other and with other educational news. Research results have the status of news, which means that some of them enter the mainstream of public knowledge and debate fairly quickly via the columns of the educational journalists. But the price that is paid for this is that research reports are not often reviewed as books, in the review columns of the papers and journals.

In this free-market situation it is a battle of all against all for attention, and there is no natural law to ensure that it is the most reputable, most serious, most significant research that gets most attention. Indeed, it is likely that there is a natural trend for that research which most easily fits in with pre-existing assumptions to catch the eye of the journalist or editor.

But reaching the news pages is only one form of dissemination, and possibly the most transitory, though the most widespread. Only a small proportion of published research receives this exposure. Perhaps more important for serious purposes are the channels of specialist publication. But here too a great many factors extrinsic to the value of the work itself have an influence on the breadth of dissemination.

The quality of writing and presentation in a research report are not necessarily a reflection of the intrinsic value of the findings (though I will argue in a moment that they are more intimately connected than is often realised). Yet there is no doubt that a poor research well-presented is as likely or more likely to achieve publication than a good research poorly presented. Publishers' editors and editors of journals cannot help being swayed by considerations of readability and presentation.

It may well appear to research workers who submit work to journals or publishers and then have long periods of waiting to endure, with no certainty of ultimate success, that the channels for publication are narrow and constricted. Yet these frustrations are no more than authors of any kind endure and have always endured in relation to publication. If there is a law of nature in the field of dissemination it is probably the law that obtained in the field from which the metaphor derives: many seeds must be cast in order for a few to reach fruition. So long as this free market situation obtains, all one can do is try to pin down the factors which make a research report, whether in the form of an article, a paper or a book, more likely to be noticed and read. In this case, readability in the broadest sense undoubtedly looms large, and I will return to this in a moment.

Need for coordination

What alternative could there be to the free market results and how desirable or dangerous would an alternative be? One feature of the research-disseminating scene which is so obvious as to escape attention most of the time is the lack of co-ordination of research results. Each research report makes its own independent claims direct to the public. A new report on a particular topic does not necessarily commence with a review and summary of all relevant previous research up to that point, though it would customarily mention at least a few precedents in the field. The basis on which research is funded by the various munificent bodies is somewhat mysterious, but does not yet, at any rate, include a systematic review of the field in question up to that point. Even those researchers who do conscientiously 'place' their work in the context of all relevant precedents may have a fairly narrow interpretation of what is relevant. Thus one may spend years becoming, and remaining, an expert in, say, methods of classroom observation without ever feeling a need to pay attention to works on, say, child psychology, sociometry or group dynamics. This is a problem of methodology. Our methods are such that more and more specialization seems to be necessary, and each field and sub-field takes on a sort of life of its own, generating its own characteristic variations of dialect, its own research result landmarks, and, usually, its own coterie of experts. From the point of view of the consumer of research results,

the teacher, LEA officer, college lecturer or student, this creates a highly fragmented climate of information.

Of course, social reality is very varied. No two schools are completely alike. Nevertheless, as more research is done, and more is reported, it becomes increasingly obvious that we have a fairly large number of researching individuals each seeking the truth in his own way and then offering it to everyone, and very few people occupied in comparing, interpreting and estimating the relative validity of different researches, and then presenting large numbers of findings in a digested and coordinated form. Funds are rarely allocated for this intermediary but highly important purpose. Dissemination is largely left to take care of itself.

To some extent this imbalance can be attributed to the attitude of universities towards research - their recognition of original published research as the criterion of higher intellectual endeavour. The dominance of this criterion creates a situation in which a great deal of research is undertaken without any real sense of conviction, urgency or involvement. The motivation to do research stems from the need to do research, not from the need to throw light on a pressing problem. Nor do I suppose anyone would like to see a situation in which research results could not make their appeal direct to the public. Any kind of censorship would be contrary to the principles of free inquiry. But an association such as this might adopt an aim of encouraging more coherence on the receiving end of educational research results. I do not know exactly how this might be done other than pressing for some funds to be devoted to the coordination of results which I have just mentioned; but I am sure that an increase in coherence would involve the emergence of a much stronger tradition of criticism in the field. Research findings may claim scientific status, but the interpretation of research findings is more of an art than a science, and it is of course the interpretation of research findings is more of an art than a science, and it is of course the interpretation that looms largest when the results come to be written up and disseminated. Research results quite quickly come to be quoted in the more discursory literature on education, but often with very little sense of critical discrimination.

Orientation to audience and orientation to problem

It seems to me that the principal problems of dissemi-
nating research findings do not lie in availability of publi-
cation channels or receptivity of the news media. There are a
reasonable number of journals of the 'learned' type circulating
in this country; and while their editorial policies are open to
criticism on one count or another, I don't think it could be
claimed that any of them suffer from an embarrassment of riches
when it comes to selection of material. At the same time, book
publishers are becoming increasingly aware of the saleability o
research reports which have some claim to general public intere
and educational journalists are fairly avid for research result
which have some claim to newsworthiness. Many criticisms might
be levelled at the judgment of publishers and editors or the
treatment of research by journalists; but there is not a seriou:
lack of publishing space available for the field of educational
research in general, relative to other equally worthy fields of
endeavour.

It is tempting to an editor to locate the problem in
writing skills. This is the phase immediately preceding the one
for which he is responsible. However, while research reports
are often poorly written, like so much of the prose with which
our society is glutted, I do not believe this is a fundamental
source of the communication problem. Bad writing is not a
mere failing: it is a creative defence. In order to write well
there must firstly be something to say that is felt to be
genuinely important and secondly a sense of audience, and a
sense of respect for the audience. One or both these things
are frequently lacking in educational research writing. And
where they are lacking, it is impossible to add them cosmeti-
cally afterwards. They can only be generated from within the
conduct of the research itself. Obscurity abounds where there
is an underlying fear of revealing a poverty of ideas or
motivation.

The problem goes deeper than the question of whether or
not the researcher possesses certain communication skills.
Certainly a great deal depends on the ability to write well,
to be selective, to make technicalities moderately intelligible
to the layman, and so on. Such abilities, however, are not
simply fixed attributes which one either happens to have or
not. The degree to which they are called into play and

refined upon is governed, I believe, by deeper principles which
have to do with the way in which the research itself is set up
and conducted. Orientation to audience, the underlying principle
of good writing, is closely connected to orientation to problem,
an underlying principle of good research.

The fact that dissemination is coming to be regarded,
rightly in my view, as a vital concern for educational
researchers suggests that there is a growing dissatisfaction
with the image of research as an inert mass of knowledge. If
it remains uncommunicated it tends to remain ineffective.
Behind this is the dawning realization that the audience of
educational research - administrators, teachers, parents - is
made up to a large extent of the very people whose actions
form the subject of educational research. Orientation to
problem and orientation to audience are not so much, then, the
beginning and end of a linear process as the meeting points of
a cycle. Essentially the researcher is not talking to his
sponsor about the behaviour of a third party, but to the whole
educational community about its own behaviour. Or to put him
even more firmly into his social context, the educational
community is talking to itself through, and with the help of,
the educational researcher.

The American researcher James Coleman has argued that the
reason why much educational research seems baffling or dis-
appointing to sponsors and practitioners is that researchers
tend to base their procedures on an implicit model of research
as the amassing of knowledge instead of as guiding social action.
Sponsors of research, having laid down the area to be investigated,
tended to leave the formulation of the problem in the hands of
the research team, but could later be heard to complain that the
results were irrelevant to policy-making. What the sponsor did
not realize was that what to him was a practical problem had
now been transformed into a theoretical one.

The distinction drawn by Coleman is an important one but
there is some danger of being led into a false dichotomy.
Research does consist of the amassing of knowledge, and rightly
so. However, the reason why there is a wish to amass knowledge
is to throw light on problems, and urgent problems at that; and
the reason why it is hoped to throw light on problems is, of
course, in order to solve them. Thus it is not a question of
researchers being on the wrong track when they seek to amass

knowledge but of their tending wrongly to focus on that as
the end-point of their concerns, rather than as being one phase
within a larger enterprise, namely the solving of problems.

There are several factors which tend to reinforce this
atrophied perspective, and they all involve the acceptance of
a falsely rigid demarcation line between complementary functions.
There is firstly the dichotomy of 'pure' and 'applied' research.
Historians of science have emphasised that some of the greatest
discoveries in all fields have occurred 'by accident', by
allowing the individual researcher to pursue his interests and
the problems which fascinate him, rather than by foreclosing
his mental options with a clear definition of exactly what is
expected of him. This has wrongly been taken to mean that the
exploratory researcher is not concerned with problem-solving.
Its real importance, however, I believe to be that his work
is not geared to solving pre-determined problems: but that he
is alert, indeed exceptionally alert, to the problem-solving
potential of whatever he discovers. So far from being detached
from problems, the successful 'pure' researcher is probably
exceptionally aware of them, in their full subtlety.

A second, and analogous, factor is the belief that the
researcher's function is to reveal objective fact, not to make
value judgements about fact. This again is correct as far as
it goes, but fails to see that the context of objective fact,
the reasons why it is valued, do depend on value judgements and
subjective - collectively subjective - wishes. People want to
get the most out of education; they want to feel fulfilled;
industry wants skilled workers; government wants - or ought
to want - to spend the available money to best effect. It is
for its potential contribution to the fulfilment of these
wholly non-objective wants that research is funded.

A third factor, rigidity of role, follows from this. All
jobs in a complex society are mutually interdependent, but
this is easily forgotten when one spends years of one's life
pursuing a particular function, talking mainly to others
within the same group. Researchers sometimes tend to say 'my
function begins here and ends here; I am paid to think about
this and this and not about what goes on outside these limits'.
But it is not possible to make sense of a function which is
only one phase of a much larger cycle without thinking a great
deal about the phases before and after, and about the whole

cycle. To neglect this is to assume that there is a coherent overall plan governing the various inter-dependent functions. In fact, of course, the various interdependent functions evolve in a spasmodic and unplanned way, and can only become coherent by the mutual awareness, adjustment and dialogue between all the participants.

Fourth and final of these factors reinforcing the atrophied perspective is, seeing oneself, the researcher, as responsible to one's sponsor without seeing that the sponsor is responsible to the community at large and that this, therefore, is also where one's ultimate responsibility lies. I do not mean that one should mentally ignore the sponsor but that one should be able to see and respond to the range of legitimate forces governing the sponsor's governance of oneself, the researcher.

I will conclude this very inconclusive discussion-paper with a large and perhaps provocative question. If my diagnosis is correct, one of the main impediments to effective dissemination of research results is the lack of orientation to problem-solving in the conduct of research. In order to communicate effectively with an audience in whose hands the solution to educational problems ultimately lies, one must regard the reader as a self-directing intelligence. But the audience of educational research. The prevailing modes of educational research, however, tend to treat their subjects purely as behaving organisms, and eschew the whole factor of self-directing intelligence, that is, of consciousness and reason. The resulting description of behaviour therefore tends to be one with which a reader cannot easily identify - it is an external, remote, abstracted account of something which he experiences as a complex, personal, reason-saturated activity. The question, therefore, is whether we are justified in hoping for a wide and sympathetic reception for data based on this limited methodology.

Discussion Report

The present constraints on the dissemination of research
in a free market is such as to make researchers neglect their
audience. The trend is for writing to be aimed at peers
rather than practitioners but even so the criteria for
publicising findings are more often related to fashion, reada-
bility, presentation and news worthiness rather than to their
importance to the appreciation of the phenomena being studied.

One critical element in educational research is that the
audiences is also the population which implies a greater need
for the data to be communicated in terms readily comprehensible
to the users. The problems can be summarized under four main
headings:

a) the apparent dichotomy between pure and applied research
b) the difficulties of establishing "value free" research,
c) the restrictions of the researcher's role being too
   narrowly defined,
d) the need for the researcher to extend his responsibility
   beyond his sponsor to society at large.

The researcher tends to return his data to his audience
in ways that have changed their reality for them. This can
only be justified if it gives new insights to the audience.

The communication process is an interactive one and one
problem of present practice is that it assumes a relatively
homogeneous, passive audience. It would appear that the
active involvement of practitioners in the processes of
research leads to more positive attitudes to research. The
preoccupation with communication using the written word
reduces the amount of active participation and it is important
to explore other forms of communication, including discussion
forums and workshop.

A particular need is for mediating agent between the
researcher and the audience. One function would be to inter-
pret data and to select what sort of research material should
be disseminated. Practitioners probably lack the time to
study the minutiae of research design but are thus exposed
to taking findings on trust.

The problem of why research needs to be communicated
would have to be resolved and a major function would be to
establish a climate among practitioners that is receptive to
research as a way of guiding action. This is probably more

important than the function of "reproducting researchers".
This would lead to greater feed-back from the practitioners
to the researchers. Eventually, it will be essential to
explore ways in which practitioners can be more involved in
research activities and researchers in practice.

The role of BERA in resolving such problems is promising
but confused. A journal that only repeated the practice of
current journal that only repeated the practice of current
journals in which papers were written for fellow researchers
was not seen to be helpful, but a regular "occasional"
publication, designed to present multi-disciplinary critiques
of educational researchers, on particular themes, would serve
a useful purpose. Crossing the disciplines would reduce
specialised jargon and develop further insights into the
problems. The publication could also include articles review-
ing the state of the field of research in given areas. In
addition, a rotating editorship would help shift perspectives.
A further role for BERA would be to promote workshops and
conferences for both research workers and the general practi-
tioners (in the widest possible sense to include parents and
other interested persons). The National Child Bureau has set
up successful meetings of this kind to disseminate the results
of their research.

SYSTEMATIC OBSERVATION IN NATURALISTIC SETTINGS

DONALD McINTYRE

First I shall try to define what I mean by 'systematic
observation in naturalistic settings'; next I shall indicate
some of the purposes which I believe can best be achieved in
this way; then make explicit some of the assumptions implicit
in such as approach; and finally I shall mention some of the
major problems which have not yet to be resolved.

Definition

By 'systematic observation' I mean that the observer
focusses his attention upon prespecified aspects of class-
room activities and describes these according to a prespeci-
fied system. In principle, and to a large extent in practice,
there need be no limit to the nature, the breadth or the
number of aspects upon which one focusses attention - that is
a function of one's concerns, one's manpower and one's
technology; what is crucial is that these aspects are pre-
specified. It is the prespecification again which is crucial
with regard to the descriptive system used: one might use a
category system (e.g. how much of the time was spent in
teacher talk) a sign system (e.g. how often did a change in
the social grouping of the class occur) or a rating system
(e.g. how clear were the teacher's explanations): all of
these would be systematic so long as the way in which they
were to be used had been decided upon before observing the
lesson.

By 'naturalistic settings', I mean that the teacher and
pupils are not constrained to do anything other than what
they would be doing if they were not being observed. The
setting is naturalistic in so far as the teacher and class
who are meeting, the time and place they are meeting, the
content and purposes of the activities they pursue, and the
methods by which they pursue them, are not controlled or
deliberately influenced by the researcher. There is of
course a continuum from naturalistic to controlled settings;
and because the very presence of the observer or of recording

equipment may influence the behaviour of teachers and pupils, it is not possible to draw any clear dividing line. But broadly speaking, the setting is naturalistic if the activities are part of the normal ongoing work of the school. I would point out, however, that this does not exclude experimental studies, if the experimental variables are included not solely for the purposes of the research but as practical attempts to change the ongoing pattern of activity - such variables as inservice training programmes, different bases for pupil allocation to classes, or new curriculum materials. The type of approach I am discussing, then, is one in which the activities being observed are not prestructured by the researcher, but the observational procedures and descriptive systems are prestructured.

Purposes

For what purposes is this approach useful? To give a very general answer first, I suggest that it is a necessary approach if one is seeking valid generalizations about classroom activities. The claim that systematic observation is necessary for valid generalization to be possible is based upon the evident complexity and, multi-faceted nature of classroom activity. To observe a classroom in an 'open-minded' way, that is, not to determine in advance what aspects one is concerned with or not to choose in advance the terms in which one is going to describe these aspects, automatically implies that one's perceptions and description of what happens in that classroom will be selective in uncontrolled and usually unconscious ways; and when one enters another classroom in an equally open-minded way, one will again be selective, but will not know to what extent the selectivity is of the same sort as in the first classroom. Furthermore, without precise operational definitions in advance, one cannot tell to what extent different words mean the same thing, or the extent to which the same words, applied to different classrooms - mean different things.

Now it is perfectly possible that non-systematic descriptions of classroom may give one more insightful accounts of what is happening in them than do systematic descriptions, or that non-systematic observation may be better attuned to identifying the different salient features of different classrooms. My point is that generalizations across classrooms

(or even generalizations across occasions) are unlikely to be valid unless one uses systematic observation. And with even more confidence, I would assert that it is only throug systematic observation that one can <u>assess</u> the validity of one's generalizations.

The advantages of naturalistic settings in allowing valid generalizations appear even more obvious. If the researcher manipulates the context or the behaviour of teacher or pupils, in any way, then we have to say that as yet we simply do not know how far his conclusions can be generalized to 'natural' settings. (It should be noted, however, that there is nothing inherently different between the ways in which researchers control or manipulate class-room settings or behaviour and the ways in which these are manipulated and controlled by a multitude of administrative, political and other factors. So that just as we cannot generalize from controlled or laboratory settings to 'natura settings, we cannot generalize from one 'natural' setting to another. In order to assess the generalizability of any conclusions from research in naturalistic settings, it seems necessary for researchers to identify their sample of classes in terms of pupil, task, architectural and organiz-ational characteristics. That this has not generally been done is not surprising, when one considers the difficulty of conceptualizing and measuring some of these variables, and the fact that, as yet, we do not know which of these many variables are important.)

Returning to the purposes for which I believe systematic observation in naturalistic settings to be appropriate, I have argued that it is a necessary approach if one is aiming to generalize about classroom behaviour, with the qualificatic that it is important to attend to the range over which one's generalizations may be valid.

More specifically, there are four main kinds of question for which I think this approach is desirable:

(1) To test the validity of descriptive models of classroom activity. (e.g. Calderhead and Morrison, 1973).

(2) To identify the effects of teachers' and pupils' classroom behaviour upon pupils' subsequent achievements, attitudes, self-concepts, and behaviour. We need, I believe, to consider a wider range of 'product' variables, and also to

specify these variables more precisely than has generally been the case.

(3) To assess the effects of attempts to influence teaching trough pre-service or in-service training or through the introduction of new curriculum materials. This can profitably be combined with process-product research by, for example, asking how the different ways in which teachers implement curriculum innovations are related to the effects which they have upon their pupils' attainments and attitudes.

(4) To identify the effects of contextual variables of the various kinds already suggested on the relationships between classroom processes and outcomes and on the effectiveness of attempts at innovation.

Implied in what I have said, then, is the suggestion that several kinds of relationship can and should be investigated concurrently, thus involving intensive use of the data collected through systematic observation in classrooms.

Assumptions

What assumptions are implicit in this approach to classroom study? I think there are several which might well be questioned:

(a) Most obviously, one assumes that generalization is possible and useful, generalization for example about how classrooms will be affected by the characteristics of the pupils and the organizational setting, generalizations too about the patterns of teacher behaviour most likely to have various effects upon pupils. Although some teachers appear to doubt the validity of such assumptions, they appear to be necessary assumptions if research on classrooms, if teacher education, or indeed, if any educational policy at all is to be worthwhile.

(b) A related assumption is that the abstract, predetermined categories in terms of which we describe classroom behaviour are sufficiently appropriate for describing the behaviour that they do not seriously distort the events which they purport to describe. David Hargreaves has in his book quoted Cogan on this point: 'The truth is that these data are so attenuated, they are so remote from the sights, sounds, the smell, the feel, and the sense of the classroom that the reality escapes us.' This complaint may of course be made against a particular observations system or against all systematic observation; if the latter, it appears to me either to be implying that valid generalisation is impossible or that my next assumption is false.

(c) This is an assumption which applies to most
observation systems except those which involve
only ratings of the whole lesson or large parts
of it: it is that one can usefully describe what
is happening on a classroom in terms of observable
specifiable behaviour without trying to find out
the 'meaning' which such behaviour has come to
have for the teacher and pupils in that particular
classroom. I expect David will have more to say
about that assumption.

(d) A fourth assumption is that a deliberately partial
description of classroom activities is not in
itself distorting. Does it not distort more than
clarify to describe one aspect of classroom activit
and to ignore the several other aspects which may
covary with this aspect? Our assumption must be
that it does not or, I would prefer to say, that it
need not.

(e) A particularly important assumption, which is
implicit in those I have already mentioned, is that
one can validly categorize ways in which language
is being used. I suspect that this is a valid
assumption in principle but that we may need to
become a good deal more sophisticated before we
can make it in practice.

(f) Finally - though this is not, I fear, a compre-
hensive list of assumptions - let me quote
Rosenshine: 'It is possible that studying teaching
in natural settings is unproductive because the
settings are not functional for the desired out-
comes.'

The assumption which we make, then, is that the settings
which we are currently faced with are settings which are both
potentially productive and likely to be little changes in
the forseeable future. This assumption is, of course,
challenged both by de-schoolers and by educational technolo-
gists. It must be confessed that it is a fundamentally con-
servative assumption. On the other hand, it is an assumption
which is equally implicit in other kinds of naturalistic
research. And it might be claimed in favour of the viewpoint
which Gordon will be putting forward that the long-term pay-
off is likely to be higher from research which does not
assume the kinds of natural settings which we find in the
schools today.

Problems

Even if we are prepared to make these various assumptions,
there are more problems in adopting this research approach
than, I think, are commonly realized. Let me mention four of
these briefly.

(1) <u>Choice of variables</u>: Systematic observation can be used only after one has decided <u>what</u> is worth observing and <u>in what terms</u> it should be described. How then should we set about identifying worthwhile variables or categories to be observed?

I would suggest (a) that we need to rely much less on personal hunches.

(b) that we should be much more active in translating psychological and sociological theories into hypotheses relevant to classrooms.

(c) that we should do much to maximise the match between our variables and the constructs which teachers and pupils themselves use.

(2) <u>Evaluate and Emotive Connotations of Observational Constructs</u>: A very large proportion of the constructs we use in observational systems clearly imply evaluation. This is most strikingly so when we <u>rate</u> teachers, for 'warmth' or 'enthusiasm', etc. But it is also the case with many category systems, such as that of Flanders.

I suggest that, for research purposes, we should be attempting to define constructs which do not have such evaluative connotations - though no doubt we shall never entirely succeed. My reasons are

(a) our data is certainly distorted by halo effects.

(b) we use such evaluative connotations to avoid the need to say what we mean. In rating 'business-like' behaviour, for example, we need to make judgements of what behaviour is <u>appropriate</u>. We may be able to make such ratings reliably, but any conclusions we reach about the importance of businesslike behaviour will be realtively useless unless we can make our <u>criteria</u> of appropriateness explicit.

(3) <u>Definition of Units</u>: How should we carve up the flow of classroom activity into bits? What units can we use?

Time units are the most obvious, but they pose some serious difficulties. Very often they appear to lack phenomenal relevance; and unless the time unit

may have no uniformity.

Other units can, however, be equally arbitrary, and often lack neutrality in the sense that they cannot be identified without also being classified. It is in our definition of units, even more than in our classification of them, that we often seriously lack either a theoretical or a phenomenological rationale.

(3) Statistical models: We need to cope with multiple codings, with chains of events, with non-linear relationships, and with the possibility that several different patterns may produce the same effects. These are just a few of the factors which make our conventional correlational models quite inadequate for coping with the richness of the data which we can collect and should be collecting. The fashionable decrying of quantitative or psycho-metric analysis of classroom activity is to some extent justified: but it is justified not because, as often claimed, sophisticated statistical techniques distort the reality of the classroom; on the contrary, it is the lack of sophistication of our statistical techniques which is potentially distorting.

SYMBOLIC INTERACTIONIST PERSPECTIVE ON CLASSROOM STUDIES

DAVID HARGREAVES

1.   There is perhaps an important moral to be learnt from the
history of learning theory applied to education. Rather than study
children in their 'natural' learning environments (which would
include the classroom), we applied the concepts and accepted the
presuppositions of those psychologists who were studying learning
by relatively simple organisms in laboratory situations. The result,
A near disaster for educational psychology.

2.   In classroom studies, the owrk of Ned Flanders typifies what I
would regard as a mistaken and fruitless approach to the understanding
of classroom life. Among the main objections to such an approach to
interaction analysis are:

   (a) the obsessive concern with measurement;

   (b) its seduction into the prescriptive teacher-effectiveness
       studies;

   (c) which is perhaps the most important objection, the neglect of
       the meaning of events to the participants. The predetermined
       categories rest on some common-sense understanding of life in
       the classroom which for the most part is simply taken-for-
       granted by the researchers and thus remains unexplicated. What
       the researcher 'sees' through his system is systematically
       distorted by the filter of the imposed and pre-constituted
       meanings and interpretations that make up the categories.

3.   An alternative approach is available in the perspectives of symbolic
interactionism and phenomenology. The first imperative of these
perspectives is to look at the phenomenon under study and to be faithful
to it. This involves an attempt to grasp the meaning of classroom events
as the members themselves (teachers and pupils) apprehend them. Among the
important questions to be asked here are :

   (a) How do the members typify events, actions and persons in
       the classroom?

   (b) What constitutes their common-sense knowledge of events
       and persons in the classroom?

   (c) What recipes (Schutz) do the members use to interpret
       events in the classroom?

   (d) What recipes do members follow to construct their actions
       in the classroom?

4.   These questions do not, and cannot, get an answer from a
Flanders-type approach. We have as yet little systematic research
in the suggested alternative perspective, but the 'flavour' of such
an approach has been captured in the 'popular' writings of Holt,
Jackson, Henry. Why are such writers much more attractive to
education students than the interaction analysis of the Flanders
type? I suggest that it is because, through looking and listening
hard, they have captured something of quality and spirit of life in
the classroom. At present, this work is methodologically
unsophisticated. But perhaps we should be prepared to take massive
methodological risks in our scientific search - or is methodology
our master rather than our servant?

5.   If educational psyhcologists had had the courage to make a
direct study of learning in school, then they might have generated a
theory of learning that would have put the mandarins of psychology
departments, lost in their observations of rats and pigeons, utterly
to shame. Those who study classroom interaction could, if they would,
generate a theory of interaction at both the substantive and formal
levels. But I fear we shall be destined to be mere technicians for
Flanders or pale shadows of those who make theoretical and methodological
advances elsewhere.

APPLICATIONS OF LANGUAGE THEORY TO STUDIES OF CLASSROOM BEHAVIOUR

W.J. Maynard, Dundee College

If we accept that teaching consists of a series of acts of communication, and that these acts are predominantly verbal, then we must accept that the spoken language must figure prominently in any description of classroom behaviour. Studies involving language may usefully be thought of in terms of the generality of their aims and methods, whence derive their scope for achievement and their limitations.

Level One may be designated 'the Macrocosmic', since it attempts to take in the whole language output of the classroom over sustained periods. From this potentially vast amount of material, it seeks to provide a descriptive system for language behaviour which will be applicable to both teacher and pupil over the whole range of their possible classroom interactions. Because of the quantity of the potential data and the complexity of language itself, it needs specialist observer-recorders, highly trained in a specific system of language-function analysis. Even so, the sophisticated detail of such systems preclude instantaneous recording, and assessments are thus based on either written or taped transcripts. It may prove that the best of such systems, such as Bellack's, or that of Sinclair and his Birmingham team, may help to establish regular patterns of language exchange, strategies of language use, which represent the normal teaching approaches of a particular individual, or even of a large body of teachers. Some suspicion, however, must always attach to systems requiring an analysis at second-hand of the purpose of an utterance, part of whose meaning must lie in the classroom context. Moreover, it may be argued that teaching is not a 'game', in the sense intended by Wittgenstein, and that no single body of rules can therefore be said to attach to it. Rather we would need multiple sets of rules relating to the roles played by teacher and pupils in the course of the many highly varied activities which go to make up the life of the classroom. Finally, the systems tend to assume a simple serial relationship among the verbal acts of a 'teaching cycle' or 'interchange'. (Fig. 1). In fact, while a lesson may have

a main flow axis, class participation will often take the
lesson away from this axis on to other planes, or return to
an earlier point of the main flow. (Figs. 2 and 3). Any
general description of classroom discourse should have some
means of indicating the nature of such flows.

The unmanageability of the language data handled by macro-
cosmic approaches has led to what I will define, with apologies,
as a 'mesocosmic' one. Such approaches filter out some of the
linguistic or functional data in order to focus attention on
certain specified broad areas of language use. These areas are
defined by the research interest, such as O. Smith's 'logical
dimensions of language', or Gallagher, Ascher and Shaffer's
'cognitive-directed behaviour', or even Flander's concern with
what he chooses to call 'direct' and 'indirect' methods of
teaching. Such researchers still offer a set of functional
categories, but it is now a limited and readily assimilable one,
reflecting the narrower range of the research objectives.
Observers are therefore more readily trained. However, the
allocation of language to appropriate functional categories
can still involve some very hazardous judgments, especially,
as when with Flanders and the systems deriving from him, the
recorders are asked to force all the language used into the
same set of semi-arbitrary categories.

At the third, or 'microcosmic' level, we are dealing for
the first time with the raw data of language, the actual words
used instead of speech filtered through a system of functional
analysis, which is often quite non-linguistic in conception.
The researcher achieves this concentration by very specific
delimitation of the situational context of the utterances in
which he is interested. One could say that the whole situational
component of utterance has been converted into a set of experi-
mental controls. To do this effectively, the researcher needs
comprehensive and reliable models of the dimensions of classroom
behaviour, from which may be located potential (in Bernstein's
phrase) primary critical socializing contexts. Figures 4 and 5
provide Guilford-like three-dimensional models for the analysis
of task-directed and socio-emotional language acts, respectively.

In Fig. 4 the shaded areas labelled A, B, C represent respectively:

A.   an attempt to <u>analyse</u> the significance of a <u>datum</u> related to the acquisition of a skill; e.g. grammatical analysis;

B.   evaluation of a concept embodied in the material content of a lesson, e.g. discussion of the effectiveness of rotation of crops as an ecological measure;

C.   information about a generalised concept directed towards improvement in a skill, e.g. information on the methods of historical enquiry.

Given such a precise definition of possible uses (and it is my belief that in the classroom management sphere, they can be made still more precise), the researcher has patiently to wait through perhaps many lessons and many months before he will accrue enough examples of the use in which he is interested on which to base any generalisations. The material he collects, however, should be precise and appropriate, since the narrowness of his recording task will enable him to pay full attention to the general purposes of the teacher and the class. I could suggest five angles from which the data might be approached:

1.   By an analysis of collocation, i.e. the kinds of words which tend to occur together in an individual's speech.

2.   By a determination of the defining grammatical features of individual or group speech, e.g. use of personal pronouns, mood, tense, impersonals, passive forms, subordinations, conjunctions and relatives.

3.   By a determination of the defining social features of individual or group speech, e.g. naming conventions, sympathetic circularities ('You know', 'see', 'like'), idiosyncratic personal markers.

4.   By an analysis of the 'prime focus of utterance', as suggested by Robinson, after Joos. This would determine to which of the six major features of the context of situation a particular utterance was addressed. These features have been categorized as:

        i.   Referential environment.

        ii.   Addresser's condition.

        iii. Addressee's condition.

        iv.   Addresser-addressee relationship.

    v.   Group feeling.

   vi. Medium of expression.   (See Fig. 6).

5.   By an analysis of the content of utterance into a set of appropriate socio-semantic options along the lines of the system suggested by Halliday. This would allow the location of subsidiary foci as well as the main focus encompassed by the previous method.

For anyone concerned with teacher education, these approaches have obvious practical merits. They allow the attention of the student teacher to be focussed on specific details of his verbal and para-verbal behaviour, and allow him to suggest and experiment with his own variations of this form of delivery. Vague advice, like, 'Be more explicit', or 'Be less formal in your relations with children' can be both quantified and qualified in terms of specific linguistic and para-linguistic features of utterance. At the same time, however, such approaches will tell you nothing of the framework or flow of discourse, which would still have to be investigated by means of one or more of the macro- or meso-cosmic methodologies previously considered. There has always been more than one way to truth.

NON-NATURALISTIC STUDIES OF TEACHING

GORDON MacLEOD

The argument to be advanced in this paper is that the study of teaching in a non-naturalistic context is both desirable and necessary.

Naturalistic classroom research may be thought of as the study of teaching as it exists; non-naturalistic classroom research may be thought of as the study of teaching where control of the independent variables is granted to the investigator. The purpose of this latter is to introduce much-needed rigour to the study of teaching processes in order to maximise internal validity, without, at the same time, jeopardising external validity. The situation to which this paper is addressed was concisely outlined by David Hughes in the American Educational Research Journal:

> research on teaching may be fairly dichotomised into studies which have been rigorously controlled and studies which have been carried out in school settings.

Although research on teaching is faced with all the conventional control problems of psychological research (see, for example, Heath and Neilson), there do seem to be additional difficulties for the process-product researcher faced with the task of evaluating "effective" teaching.

This paper will outline three such problem-areas-coverage of curricular content, control of the teaching behaviour itself, selection of pupils - and will examine three non-naturalistic studies which have gone some way to solving such problems.

Coverage of Content

In any process-product study, correlational or experimental, it is important that the groups of pupils being compared should have been exposed in equal measure to the content being assessed or sampled by the criterion test(s). Similarly, any criterion test must sample not only the pupils' coverage of content, but must also sample the kinds

and levels of knowledge and comprehension which the teacher
is attempting to teach.

Only when such variables come under the investigator's
control, does an effective strategy seem possible. This
control is exemplified in a study of lecturers' explaining
effectiveness by Gage and his associates in which the
researchers specified the instructional content by providing
the lesson materials and explicitly specified the objectives
both by description and by the provision of examples of the
criterion materials. Because of these constraints and
because the instructional period was short, content-analyses
of the recorded lessons became feasible and indeed a signifi-
cant relationship was found between the categorised content
and one of the criterion measures.

Control of Teaching Behaviour

The most immediately apparent way for the researcher to
gain control of teaching behaviour is through the 'true'
manipulative experiment, in which groups of teachers are
trained to behave in prespecified ways and comparisons of the
differential effects and effectiveness are carried out. This
design is, however, fraught with difficulties not the least
of which is that of ensuring that one's training procedures
do lead to change, and the results of such studies as there
have been, have been generally disappointing.

Other methods of control do exist, however, and probably
the most extreme of these have been by Church and Hughes at
the University of Canterbury in New Zealand. In these studies
the control taken by the investigator was either of the
rule-following kind or by the use of scripts, learned word-
for-word by the teacher. These studies come near to exercising
the degree of control usually associated with laboratory research
whilst retaining their generalisability to classroom instruc-
tion, and do seem to represent a possible new strategy for
control of the teaching situation. As Rosenshine and Furst
note:

> it is the design ... rather than the particular results
> which merit the most attention for future research on
> teaching.

Control of Pupil Ability

In most naturalistic settings it is difficult if not
impossible to control adequately for the initial ability and
knowledge of the pupils, and, as a consequence, most researchers
have accepted this constraint and have attempted to use indirect
or statistical control of these variables rather than direct
or experimental control. The most usual procedure has been
to give a pre-test of knowledge and/or ability and to 'correct'
or adjust the post-test scores on the basis of these. Tech-
niques used have ranged from the notoriously error-laden gain
or shift scores to the use of complex analysis of covariance.
The dangers in the use of gain scores are well-known; the
difficulties inherent in the apparent sophistication of analysis
of covariance are less well-known; the difficulties inherent
in the apparent sophistication of analysis of covariance are
less well-known. Firstly, as Cronbach and Furby point out,
the use of residual scores to adjust for initial differences
in samples is very much in question if these initial samples
have not themselves been randomly drawn, or if these samples
have not been selected on the basis of their pre-test scores.
Secondly, as Lindquist points out, a further assumption for
analysis of covariance is that the 'pre-test' scores are them-
selves unaffected by or uncorrelated with the treatments, both
of these being unlikely assumptions where pre-tests are
designed to measure pupils' abilities in the content area in
which they are to be taught. Thirdly, the acceptance of this
latter assumption in analysis of covariance means the neglect
of possible interaction effects among the pre-test measures
and the teaching behaviours being studies. Fourthly, when
analysis of covariance is used to adjust for large initial
differences, there arises the problem of interpreting the
meaningfulness of the resulted adjusted scores.

According to Flanders,

the ideal experimental design, which is usually beyond
the reach of the researcher, would involve the random
assignment of pupils to classes from strata based on
ability.

It is, however, only in naturalistic contexts that this
'ideal' design is outwith the reach of the researcher, for,

in studies at Stirling University by myself and Donald MacLennan, advantage has been taken of the microteaching facilities to allow rather more control of the teaching situation than is usual. In these studies groups of student-teachers in three subject-areas, English, History and Biology were asked to teach specified subject-matter, with specified aims and objectives and with some pre-knowledge of the pupil test materials, to 'micro-classes' of five pupils. All the lessons were video-recorded, typescripts were made from the video-recordings, and from the typescripts the teachers' questioning and reacting behaviours were coded into over twenty categories and the frequencies of these categories were then correlated with a selection of pupil outcome measures, including the results of achievement and attitude tests, and measures of pupil behaviour during the lessons.

The important advantage provided by the use of classes of five was in being able to follow Flanders' recommended ideal design - 'the random assignment of pupils to classes from strata based on ability'.

As a check on this method of control of initial ability, the correlation was calculated between the History and English achievement test scores of thirty classes who had been taught both lessons. The product-moment coefficient between these two sets of scores was found to be .07, suggesting that random assignment from strata was very successful in diminishing the amount of variance in the test scores accounted for by initial ability.

Whilst the analysis of the data from these studies is not yet complete, several of the preliminary results have been very encouraging. In the English lessons, for example, it has been found that 73% of the variance in the achievement test scores can be accounted for by only two of the composite predictor variables i.e. a multiple R of .85. Similar high multiple correlations seem to be emerging for all thirteen of the criterion variables.

The three sets of studies referred to seem to demonstrate attempts to determine internal validity through the use of non-naturalistic research strategies. Whilst it is important that research designs should involve the maximisation of both

internal and external validity, it is the former which is 'basic minimum without which any experiment is uninterpretable' (Campbell D.T.). The object of the approach being advocated and exemplified is to maximise the internal validity of classroom researches whilst at the same time not jeopardising external validity. There seems no single solution to maximising both kinds of validity, yet it is important, at a time when educational research is subject to pressures for 'applicability' and 'generalisability' to bear in mind that internal validity is the 'sine-qua-non' of experimentation.

Discussion Report

It would be quite impossible to summarize the papers presented as they covered a wide range of methodological perspectives, including the most open-ended of symbolic-interactionist pleas and the most carefully controlled structuring of classrooms as laboratories. For this reason, the participants presented their papers separately. There was a similiar range of perspectives within the audience and so the discussion ranged back and forward between methodologies and values, applications and theories, and philosophies and techniques. To summarize it can only fail to do justice to the subtleties of the arguments and the interactions and it epitomizes the dilemma of the classroom observer about how far he can impose his reality on the phenomena and yet still make it meaningful and real to his audience. So if there is lack of agreement between this account of what happened and all the other participant observers, comfort may be sought in accepting that the reporter's paradigm is individualistic and impressionistic and serves only to highlight the need to keep all the options open in deciding on appropriate methodologies for studying classroom behaviours. Certainly it was felt that any attempt to establish a hierarchy of best (or better) techniques would be to deny the range of insights necessary to further understanding of the realities of teachers' and pupils' classroom behaviours, of what "really makes them tick".

The first issue was essentially related to the degree of reality or spuriousness of controls. In a discussion of the effects of "praise" on pupil achievement, it was argued that observation would serve to identify a hypothesis, but that controls would further the testing of such a hypothesis. No agreement was reached.

The second issue was concerned with the degree of common sense in psychometric testing. It was posited that, for all its alleged artificiality, sound psychological research did give insight into classroom practice and that the plea for "common-sense" was itself misleading. Classroom behaviour could be explained in terms of sympathetic ritual magic, itself a long way from manifest common-sense, and this could also be beneficial. It was concluded, for some, that direct

contact with classroom phenomena was not the only way to
gain accurate analyses of how teachers taught and pupils
learned.

This lead to a discussion of the major dilemma for
educational researchers that teaching requires action.
Decisions have to be made about methods and approaches
and so both researchers and teachers have to come down in
favour of one approach or another. If the researcher goes
for controlled approaches, he has to meet the demands for
increasingly complex and sophisticated statistical techniques
which tend to take him further and further away from the
reality of the individual in the classroom. If he goes for
a much more direct, action-orientated approach, the phenomena
are changed by the developing research strategies. This
lead to a general acceptance that there _is_ an unresolved
dilemma : that although teaching is different from research,
research into teaching involves the practice of teaching.

This brought up the issue of ethics and there was
considerable agreement about the overall objectives here,
whether it was related to teacher training, curriculum
development or research intervention. The common goal was
the _autonomy_ of the teacher and the common strategy to
offer data to help a teacher develop an informed critical
awareness. But the different protagonists were either so
sure of the rightness of their perceptions and methodologies,
or the wrongness of those of their opponents, that they were
convinced that a different set of values was being described.
If an example of selective definition of situations needed
presenting, the ensuing discussion probably gave it. From
the reporter's viewpoint, it seemed that the accounts offered
by the phenomenologists seemed less just than those offered
by the psychometricians but that may merely be an attempt
to offset personal bias.

Nevertheless there was a major anxiety about the ethics
of "arranging" teaching situations, e.g. by using predetermined
scripts, and of the conceptions of teachers and pupils implied
and whereas not everyone felt that this did violence to the
individual teacher's autonomy, most felt it was problematic
to prescribe for personal relationships.

Some attempts were made to identify the culprits to whom the blame for such moral dilemmas should be ascribed, the researcher or the consumer, but as with the sociological debate between action-theorist and structural-functionalist, it was held that the real issue lay in the individual's value system and his view on the nature of man. The moral responsibilities lie with all parties involved. Those who initiate research, those who choose research methods, those who make decisions about its use, and those who use it.

This led to the pertinent question about the use to which research data, irrespective of the validity of their sources, are put. Is research to add to the sum total of knowledge; is it to guide teachers, or teacher trainers; or is it for policy making? Only in one sense was the question answered and that was to suggest that it depended on the researcher's own values. Several values were aired, mostly in relation to training : e.g. to help students form their own analyses of teaching; to test the assumptions of their training; to choose from techniques offered; and to improve their initial competence. The most fruitful area was exploratory and hinted at the idea that research was necessary to explore the relationship between the teaching function and the research function in order to generate an adequate theory of teaching. Implicit in this is the concept that an educational research theory can be generated out of the teaching process rather than be borrowed from related areas of study.

A touch of iconoclastic reality was presented in a contribution which deplored any analysis of teaching that was not in the interests of pupils. It was suggested that too much teaching and research was mainly to promote the interests of teachers and researchers and that any approach that encouraged "objectivity" in teachers, "even from studying misguided research" was necessary for better teaching.

The subsequent discussion of teacher education was interpreted as flight from the more fundamental issue of the assumptions that lay behing the choice of research methodologies but this neglected the reality of the experiences that shaped all the participants' contributions. For almost every member of the audience, research values are made manifest by their involvement in training. This became more explicit in a discussion of the reciprocal nature of role ralationships and the problems of reconciling one's own identity with the expectations that make up one's role. Even more important were the organizational features that shaped this pattern of expectations as these regularly determined the conclusions an audience (the role others) would draw from any particular role player, be he researcher, trainer or teacher. Intended descriptions often became prescriptions for action because of the status of the source as, for example, when a College tutor offers a course of action. He is making explicit his expectations and students are likely to be constrained to fulfil them. Thus the dialogue between researchers and teachers over values only makes sense in an organizational setting.

In a characteristically circular way, the discussion returned to the problem of values by trying to resolve the question of whether the real function of B.E.R.A. was to get research over to teachers. The implications of what research and its relationship to policy-making stimulated the controversy of whether "systematic analysis" was better than trying to "enter into situations" to solve problems. One major issue was whether the present policy-dominated approaches to educational research were premature. Description was necessary first and systematic analysis might follow "once we can really offer explications of how teachers work", but the suggestion that theory construction could arise out of open-ended observation and hypothesis testing from systematic observation was not really acceptable. It was reported that the former did promote ready co-operation from the staff of schools, schools, but members were still asking whether the two approaches were all that different even though the protagonists were sure they were!

Finally, it was argued that the question was one of strategy rather than values. If research was not designed to extend theory, it was theological; if it was not designed to test hypotheses it was political; but if it was designed to test both it was scientific. A humanities dimension was added but that seemed to bring the discussion back to its starting point. Educational Research requires many approaches; or was it methodologies, or was it values; or was it strategies? Perhaps it was all of these! All knowledge has an element of the situational specific in it, the difficulties lie in how to generalize about it. What was universally acceptable was that there should be no hidden censorship in its dissemination, whether this arose from methods or values.

Finally, a congratulatory note came in. Members were pleased about the good tempered non-emotive level of the discussion and hoped that B.E.R.A. would continue to sustain this level of rationality. It could, however, be argued that the absence of heat reflected an absence of deep concern, so perhaps there is not even agreement about that!

This report is signed merely to re-emphasize its subjectivity.

<u>Peter Chambers</u>

CHANGING STRATEGIES FOR CURRICULUM EVALUATION

Wynne Harlen, University of Reading

There is no better proof that curriculum development is a growth area of educational research, as Professor Nisbet has said, than is provided by looking at the mushrooming publications on the subject, and especially those concerned with evaluation. Anyone surveying the literature on evaluation of curricula must be struck by the wealth of models and strategies published in the last five years compared with the number of offerings in the five years before that. There has certainly been something of an 'evaluation explosion'. At least an explosion of exposition and theorising; perhpas had this been accompanied by a similar amount of activity in the practice of evaluation then what has been written might be more helpful to those working in the field. Even in 1969, the list of publications on curriculum evaluation cited by Baker included eighty items, only six of which were empirical studies, the rest being commentaries or theoretical proposals. Since that time the imbalance between empirical and non-empirical subjects of publications has not improved.

However, this is only mentioned to introduce the situation which we find, in which there is a plethora of models for curriculum evaluation. The central concern here is to raise the questions of why there should have developed so many different views about strategy, why earlier strategies were found wanting, and what are the directions of present trends in strategies. It is hoped that by looking at the changes which have taken and are taking place and by speculating about the reasons for these changes, we may come closer to uncovering some guidelines about what curriculum evaluation should attempt to examine in order to inform the decisions which have to be taken in curriculum development.

To begin let us look at the major strategies which were used in the early 1960's when curriculum development as we know it now began in this country. There were broadly two approaches and all strategies could be comfortably put under one or other of these heading:

The 'classical' strategy. This is essentially experimental
and employs accepted research models, techniques and assumptions.
It is based in the view of evaluation as -

> *'essentially the process of determining to what extent the educational*
> *objectives are actually being realised by the programmes of curriculum*
> *and instruction. However, since educational objectives are essentially*
> *changes in human beings . . . then evaluation is the process for*
> *determining the degree to which these changes in behaviour are actually*
> *taking place.'*

This approach led to the application of methodology more approp-
riate to experimentation in the pure sciences than to the complex
process of curriculum development. It is the 'educational
science' at the extreme left of Professor Nisbet's spectrum of
research. The strategy of evaluation as 'measurment of achieve-
ment of objectives' is illustrated by Smith and Tyler's work in
the Eight Year Study (1942), by much of the evaluation of reading
materials, and work in the first wave of curriculum development
(e.g. Harlen's evaluation of the Oxford Primary Science Project).
There is the assumption in the strategy that all important out-
comes can be tested, that the 'treatment' can be defined and
prescribed, and that the important variables can be eliminated
or controlled.

The 'intuitive' strategy. By this, is meant the evaluation of
curriculum innovation by informed opinion based on informal
observation or other informally gathered information, such as
interview or discussion. It arose from the tradition which we
have in this country of educational evaluation being carried
out for many years by HMI's and local inspectors using such
methods. Inspectors look at a very wide range of processes,
concomitants, conditions and outcomes of teaching and learning.
They generally do this subjectively, without quantifying their
findings and arrive at their conclusions by bringing their
experience and their total grasp of the situation to bear.

Since this approach had been used and been accepted for a
hundred and more years, it seemed to provide a ready model for
use in evaluating new curriculum materials. It was especially
attractive to those who felt that the products of curriculum
change are too complex and broadly scattered to be encompassed
by prescribed objectives and measured by existing techniques.

Although based in part of our tradition of education in this
country the 'intuitive' approach had some support from the
writers on curriculum evaluation in the US. In 1963, Cronbach
wrote that formative evaluation should be concerned with the
desription of outcomes *in the broadest possible scale, even at the
sacrifice of superficial fairness and precision*. Many of our first
projects were evaluated using this strategy (e.g. the Nuffield
Junior Science Project).

What has happened to these two strategies in the last ten
years? It is not difficult to see the major deficiences of both;
these have been pointed out in many ways by many authors, so it
is only necessary to touch on them here. Firstly the classical
strategy is easily criticised for

(a)   Being geared to outcomes, so that it does not provide
      information about how to change the material or point
      to reasons for achievement or lack of it.

(b)   Ignoring the characteristics of the learning environment,
      giving an oversimplified view of education as a 'treatment'
      which can be controlled from outside the classroom, and thus
      providing results which are unrealistic.

(c)   Leaving out of account all outcomes which are not readily
      measurable by traditional methods.

These and other deficiencies were found as soon as the strategy
was put into practice and given a thorough trial. It has been
indeed given more than a fair trial, for even when its deficiencies
were revealed there have been those who have clung to it because
of its air of scientific respectability. In response to its
shortcomings strategies have been developed which take into
account more than the measurement of goal achievement. The most
important changes have been to include some assessment of process
as well as of product, so that one might know just what the out-
comes were the outcome of. Other changes have been towards
gathering evidence about less tangible outcomes, attitudes,
interests, skills, etc. In some cases this has been through
the opinion of the teacher or a visitor to the class. It has
also been realised that major effects may be the result of some
factors apparently unconnected with the new material, so informa-
tion about such things as the past history of the teacher and of
the pupils, is gathered, even the age and lay-out of the buildings

may be relevant . . . so the strategies are developing towards more widely based data gathering and leading to the use of techniques not unlike those employed in the informal strategies.

Secondly, the intuitive strategy was readily criticised for being

(a) unsystematic, yielding information which is subjective and incapable of being generalised,

(b) biased by sympathy, and in many cases, involvement in the production of the curriculum materials,

(c) unscientific on account of failure to define criteria and thus not subject to the accepted description in terms of validity and reliability.

Its application led to decisions which were the result, as Kerr has said of 'persuasive discourse in which each member of the tean draws on his experience and personal judgement to arrive at a consensus of opinion' (1968). In response to such criticisms the strategy, which was essentially one for looking at processes and short-term outcomes, changed to include more systematic methods: for example in classroom observation, the use of inter-action analysis; in judgement of written materials, the use of content analysis, or a formally constituted panel of experts; in judging pupils' progress, the use of diagnostic check-lists instead of the teacher's unguided opinion. In almost all these cases the changes have been towards closer definition of criteria and more systematic gathering of data. At the same time such changes have brought with them the danger of the data gathering becoming too narrowly focussed on those aspects which can be closely defined.

In a curious way, certain trends in these strategies have brought them to the same point, or perhaps overlapping, with the previously formal now being more informal than the previously intuitive, and perhaps being subject to some of the criticisms levelled earlier at the informal strategy. Whilst at the same time the previously informal strategy has led to strategies which may be too narrow in gathering data from overconcern to define criteria in objectively indetifiable terms.

The possible reasons for the changes are many and inter-connected. Obviously the changes in values in education in the last ten years must have been the cause of changes in information which is thought useful for evaluation. But there have been

changes in information which is thought useful for evaluation. But there have been changes in thinking about evaluation itself - its role is seen as to help in making decisions, to improve material being developed or to inform those who are potential users, to guide administrators, etc., not to provide answers, to prove, judge or justify. In this role there is greater concern with trying to describe, to find out what happens and to approach some understanding of the effect of introducing changes in the curriculum. Both varieties or original strategies have been found wanting when it comes to making decisions, mainly because they ignored a large part of the system which is disturbed when changes are introduced. The lesson which can be learned from this is that we should look at the system in a wide context and gradually narrow down to those aspects which are relevant to the changes being made. We have seen what happens if we look at only one part - the outcomes in the classical strategy - and have to broaden the net in order to make any useful description.

Several trends can be discerned in the changes in strategies. One is a trend away from gearing evaluation to objectives. This trend finds its extreme form in the goal-free strategy proposed recently by Scrive. By ignoring and indeed being ignorant of the objectives of the innovation the evaluator looks for all effects and outcomes without being biased by what should or should not be expected. Another trend is towards emphasis on processes rather than products. This is necessary in the short-term, since processes have to be in operation before products can be said to result from them, but it leaves the danger that we neglect the products altogether and base our curriculum innovation on assumptions that certain processes will lead to desired products. A third trend is away from statistical methods of description, which blur out the differences between one individual or one class and another, differences which may give clues to what is going on in different situations when on the surface the same curriculum is being used. Many people are now interested in adapting materials to suit the unique require- ments of individual pupils and would like to see strategies being developed which evaluate, among other thins, the adaptability of materials. At present we seem to be too concerned with averages, suiting the majority of the class without regard to the individuals.

With all these trends it is easy to see why so many different strategies are being developed, and why few existing ones ever seem to fit new requirements. There is a severe lack of any guidelines for an evaluator to follow in choosing a strategy, but these guidelines will come mainly from practical experience and not from theorising. All this points to the need for more evaluation and more publication of evaluation results.

"EVALUATING TEACHING : A MISSING ELEMENT

WITHIN COMPLEMENTARY PARADIGMS OF

CURRICULUM EVALUATION"

Dr. Richard C. Whitfield
(Cambridge University Department of Education)

Introduction

The attention given to discussions of the theory and
practice of curriculum evaluation in various publications
since the late 1960's, subsequent to the first phase of
institutionalised curriculum renewal, now makes it extremely
difficult for anyone to contribute original ideas in this
sphere. Now that the literature yet contains many examples
of actual evaluations, nor that the tensions about
methodology have been resolved - far from it. In fact the
growing disagreement among scholars in the field about
methodology now acts as an obstacle to the actual evaluation
of educational programmes. Although second generation
curriculum developers and educational administrators are
now convinced that evaluation is unavoidably part of any
curriculum model, the tensions over both its underlying
assumptions and its methodology do little to help those who
are in a position to commission, budget for and execute
evaluative studies.

This short discussion paper pleads for a detente between
those curriculum evaluators aligned strongly to the
agricultural-botany paradigm and those aligned with the
social-anthropology paradigm. In suggesting cooperation
between 'mechanists' and 'illuminators' (though it is
doubtful if the pure form of either evaluator exists - at
least in the U.K.), I shall comment upon large survey
evaluations, with particular reference to the recently
published I.E.A. study of science achievement in which
(along with many others) I was involved. I shall point out
that the specific evaluation of teachers and teaching seems
to have been omitted in considerations of curriculum
appraisal by either of the predominant methodological
paradigms, finally posing the question as to whether this

difficult and sensitive area should be avoided in any
representative evaluation plan.

Basic paradigms of evaluation

Empirical methodologies in educational research may
be crudely contrasted in terms of the agricultural-botany
and social-anthropology paradigms; Parlett and Hamilton have
more recently used this longstanding classification in relation
to curriculum evaluation, christening the latter paradigm
in this context 'illuminative evaluation'. There is no need
here to describe at length the properties of and problems
associated with each of these two paradigms; the following
list of contrasts will suffice.

		AGRICULTURAL-BOTANY PARADIGM		SOCIAL-ANTHROPOLOGY PARADIGM
A. Concerns	(i)	Prespecified achievement criteria (products)	(i)	Descriptive of many contexts (processes)
	(ii)	Curriculum treatments (experience packages)	(ii)	Dynamic uncircumscribed interaction of an instructional system in a learning milieu
	(iii)	Pre and post-testing of samples	(iii)	Case studies
	(iv)	Measurement and prediction of change	(iv)	Interpretation of fluid events
	(v)	Scientific tradition	(v)	Socio-historical tradition
B. Shortcomings	(i)	Large samples or unacceptable and artificial situational controls	(i)	Anecdotal, impressionistic, subjective
	(ii)	Divorced from reality and insensitive to local conditions	(ii)	Non-generalisable, charismatic
	(iii)	Stable treatments assumed from concensus of concern and of goals	(iii)	Rarely quantifiable
	(iv)	Blinkered to what is objectively measurable	(iv)	Unclear what is being evaluated and by what criteria
	(v)	Neglect of atypical results	(v)	Stress on the atypical

Surveying the scene in the U.K., it would seem that
institutional and financial constraints have made pursuit
of the agricultural-botany paradigm something of a pipe-
dream for those either possessing or anxious to develop the
necessary technical skills of measurement and analysis which
have a long pedigree in the literature. Such evaluators have
had to remain content with less formal modes of appraisal
which, while very useful in a limited sense, have not been
able to test the hypotheses which curriculum innovators
have had in mind with any rigour.

In this context it is all the more surprising that the
new breed of 'illuminators' should be so vociferous in their
rejection of models which have scarcely been tried in the U.K.
in favour of others whose methodological character is either
unclear or seems to add little to the widely practiced info
informal feedback within recent curriculum innovations. The
manifesto of the closed conference of illuminators held in
Cambridge in December 1972 reflects attitudes which seek to
polarise rather than to build upon multiple traditions of
scholarship. The complexity of the educational situation
requires all the insights we can muster from whatever
discipline if curricular decisions are to be made more
rationally. To reject the traditions of educational measure-
ment is to case aside a pearl of great price; this is unlikely
to be the only such pearl, but if it is used sensitively,
rather than narrowly in the strict traditions of behaviourism,
it can help us a great deal in unravelling the interactions
between inputs, processes and outputs (if I may be allowed
the use of such perjorative terms).

Once the multivariate nature of the educational context
is recognised, multivariate studies with their attendant
challenges of measurement and complexities of statistical
analysis become both unavoidable and relatively costly. But
this is the price the educational system will have to pay if
it wishes to have its judgements based upon information rather
than prejudice, tradition or mere hunch. Increasingly society
will, as it has for medicine and engineering, demand an
accountability from educational activities. This, though
presently distasteful to many educators, may mark the
beginnings of a truly professional education service.
Evaluators are therefore going to be in increasing demand since

freedom without accountability is a luxury which the
government services, both local and national, cannot afford.
We had better therefore have our strategies and skills
clarified and prepared.

The I.E.A. study of science education

One of the fundamental problems in the commissioning of
educational research is that funding bodies tend to require
answers to questions for which it is not possible, in the
time envisaged in the funding programme, to discover water-
tight answers. In order to maintain the sympathy of potential
supporters of medium or large scale research programmes, re-
searchers are then tempted to promise more than they can re
reasonably deliver. The I.E.A. studies are in my view classic
examples in which directors of research have taken on
commissions which they could hardly fulfil in order to acquire
any research at all. Responsibility for such dilemmas cannot
fairly be placed upon any one person's shoulders - a
multiplicity of interests are involved; all that is certain
is that educational research as a whole suffers if expectations
are not adequately realised.

As many of you will know, I.E.A. (The International
Association for the Evaluation of Educational Achievement) -
originally an offshoot of UNESCO - is a voluntary grant-
supported research association to which member countries are
affiliated through their national centres for educational re-
search; the N.F.E.R. and the S.C.E.R. act as the coordinating
and administrative centres for the Association's testing
programme in Britain. One of I.E.A's major objectives has
been to help to place the study of comparative education on
a more empirical basis by the quantitative description and
measurement of inputs and outputs (and, to a lesser extent,
processes) of each member country's school system. Through
such studies it was hoped, among other things, to provide
useful empirical data for decision-making in relation to the
social and educational systems of the developing countries.

To some extent these goals have been achieved. The study
of mathematics achievement in 12 countries was published in
1967, and that of science education in 19 countries last year.
Other studies relating to Literature Reading Comprehension,
English and French as foreign languages, and Civic Education

have either recently been published or are in the press. My
more specific comments are related to the science education
study with which I am most familiar as the English member on
the International Science Committee which guided the project.
This study not only illustrates the problems of large scale
international research; it has valuable insights for further
evaluative programmes of the survey, rather than experimental
tradition. I must necessarily be unfairly selective in my
comments, both in relation to the title of my paper and the
nature of this seminar, and I will leave you to follow up
further details, particularly of the results of the study,
in the international report already published. (English and
Scottish National Reports are presently being compiled).

The vision of collecting and relating data on the inputs,
processes and outcomes of whole national educational systems
using representative samples of schools, teachers and pupils
at different age levels in different subjects or disciplines
is both imaginative, attractive and ambitious - yet at the
same time daunting. To do this accurately for a single
country would be an achievement; to achieve it reliably for
many countries with differing traditions and languages would,
you may think, represent an educational researcher's dream,
or perhaps his nightmare! Yet this dream is what I.E.A. took
on board when it became clear that financial support for such
an international programme was available, political and
technical considerations making it difficult for individual
national studies to be completed prior to international
comparisons.

So the second (post-mathematics) phase of I.E.A.
studies collected data on no less than 856 variables using
about ½ million pupils, 50 thousand teachers and 10 thousand
schools. (One might wonder : 'Whither the illuminator in all
this?' Arc lights rather than candles will be required,
and so they are!) But it is one thing to code the formal
qualifications of teachers, to count the number of
laboratory technicians and so on, and yet another to measure
reliably on an international basis the interest of pupils
in science or an index of the social control in a school,
for example, with the same test items variously translated.

Intermediate in difficulty is the construction of valid
and reliable measures of cognitive and motor performance in
science based upon a common core of syllabuses and national
curricular objectives. Yet all these tasks, and many others,
have been attempted with, as you might expect, varying
degrees of success. (This, like the majority of empirical
studies, highlights the lack of suitable instruments for
measuring the range of variables in which the aware educator
is interested; to develop a suitable career framework for the
test developer, whose skills are central to the advancement
of educational research, must be a high priority, and one
which B.E.R.A. should be able to facilitate.)

By means of questionnaires (to school principals, teachers,
pupils at 3 age levels and national educational experts), pen
and paper tests for pupils and, in a few countries, tests
of pupils' laboratory skills, I.E.A. has thus gathered
extensive data on inputs and outcomes of the educational
systems of the participating countries for the academic year
1970-1971. For the practical reasons of cost and personnel,
classrooms were regrettably not directly observed, but
attempts were made through some of the teacher and pupil
questionnaires to devise reliable descriptive measures of
classroom and school transactions (process variables). Such
descriptive measures, for example of the nature of pupils'
practical work, should in my view be further explored to
enable more rigorous testing of some important hypotheses in
such 'process' areas ot take place. (Descriptive measures
may have a potential to replace or at least supplement direct
observation.) Unfortunately financial resources, after
supplementation, became exhausted, so that the rescoring of
items, the remaking of scales and further data analyses have
to date only been possible to a very limited extent; much
data thus lies on computer tape without its full potential
value being extracted.

It is important to realise that the I.E.A. surveys
are not intended to be used as an educational olympiad in
the cognitive, affective and motor domains. The overiding
endeavour has been to try to relate social, economic and
pedagogic factors to outcomes so that more rational
manipulation of inpurs and processes to enhance the achieve-
ment of a range of goals might become possible.

There is however little comfort in our report for the
curriculum designer or pedagogue, for even in science
(perhaps par excellance a school-based discipline) we find
for example in the U.K. for 14 year-olds that <u>all</u> the
teaching variables, given that pupils are in fact studying
some science, contribute to the extent of only 3 or 4% of the
between-schools variation in achievement. This is some
twenty times smaller than that accounted for by a cluster of
background, non-school variables which are termed the 'school
handicap score' - a measure of the socio-educational ecology
of the school intake. Here there are some interesting
international comparisons, indicated by the annotated table
derived from the multiple regression analyses.

SCIENCE ACHIEVEMENT AT 14+

% variance accounted for by

Country	School Handicap Score	Learning conditions	Sex
Australia	35	15	13
England	64	4	2
W. Germany	46	26	6
Hungary	20	7	0
Japan	42	10	1
New Zealand	43	20	12
Scotland	80	3	0
Sweden	8	44	5
U.S.A.	62	11	3

One of the most difficult problems for educational research in Britain illustrated forcefully by this table is the extent to which the apparent effects of the numerous 'learning conditions' which are potentially manipulable within our educational system (curriculum, method, teachers, laboratory staff, equipment etc.) are swamped by the massive ecological contribution to the variation in scores. In other words the process variables in which the curriculum evaluator is centrally interested do not exhibit outstanding significance when the input variables are also properly considered. When in addition one considers the potential errors of measurement in the assessment of the complex array of pedagogical variables it is not surprising that the I.E.A. studies, at least in Britain, suggest relatively little to assist the curriculum innovator or classroom teacher in their pedagogical decision-making. If the effects of curriculum variables (that is a whole gamut of teaching/learning variables, not just a curriculum package) are to be teased out more strikingly, if they are not to be dwarfed by other determinants of achievement, matching studies (in which the major ecological variables are equalised between two or more groups) and/or the assessment of completely new variables become necessary.

This leads me to perhaps the most important section of
my paper in which I wish to point to an important set of
process variables which have been omitted in the I.E.A. studies
and in the strategies of the two polarised paradigms of evalua-
tion to which I have referred.

The evaluation of teaching

There has been a tendency, particularly in this country,
to reject, or at least to ignore even in the world of teacher
training, the extensive literature on teaching effectiveness.
While much of the research carried out on the qualities of
teachers which appear to be related to desired effects on
pupils is either inconclusive or methodologically suspect,
an appraisal of the literature does begin to indicate some
positive messages (see, among other sources, Rosenshine's
relatively recent review). If, either as agronomists or
illuminators, we agree that we require sensitive descriptions
of teacher-pupil interaction as a part of our studies of
curricular processes, it would appear that we cannot avoid
collecting information upon a range of relevant teacher
characteristics (not restricted merely to verbal behaviour)
in the interactive situation of the classroom.

The literature on teaching effectiveness would seem to
suggest that the following groups of teacher characteristics
are relevant to much pupil growth, and therefore for curriculum
evaluation.

1. Warmth, empathy and affiliative drive towards pupils;
    sensitivity in the use of pupils' ideas (as opposed
    to egocentric).

2. Cognitive flexibility and an ability to create a
    variety of classroom options, materials and activites;
    imaginative (rather than routine).

3. Demanding of pupils in terms of their attention and
    resource for learning; businesslike and orderly
   (as opposed to slipshod).

Each of these broad traits have an associated range of
classroom acts by teachers. It is not beyond our wit to
monitor these using observation schedules and rating scales
with a limited number of categories, supplemented by measures
of self-concept, attitude and belief (for example the M.T.A.I.,
and triadic elicitation procedures derived from George Kelly's

theory of personal constructs). Such multiple procedures can go far below the surface structure of interactive behaviour which the illuminative evaluators fear is a consequence of using codified observation.

To contemplate such measures of personal characteristics of teachers is to be prepared to use the insights of perceptual and phenomenological psychology which have been directly extrapolated to the context of teaching and other 'helping professions' by scholars such as Carl Rogers and Arthur Combs. This will, hopefully, mitigate the fears of the illuminators over the potentially excessive influence of behaviourism.

It may be that multivariate studies which incorporate teacher measures such as those which I am suggesting will indicate that these personal, individual, idiosyncratic variables are far more important to pupils than any curriculum package or school resource however skilfully compiled and designed. This is I suspect a major hunch of the illuminators, which I share. If this hunch is demonstrated to be a reality in a number of rigorous studies, teacher training through curriculum development will have to be much more sturdily complemented by teacher training through inter-personal awareness and development.

Educators, it seems to me, must each be concerned about (i) the inputs of the educational system, (ii) the nature and quality of curriculum processes and human encounters within it, and (iii) the outcomes, both long and short-term, of the engagement. Evaluation must therefore comprehensively consider both processes and products in relation to dynamic goals for the inputs as indicated in the chart.

Process/Product Evaluation of Original and Modified Goals in relation to Inputs
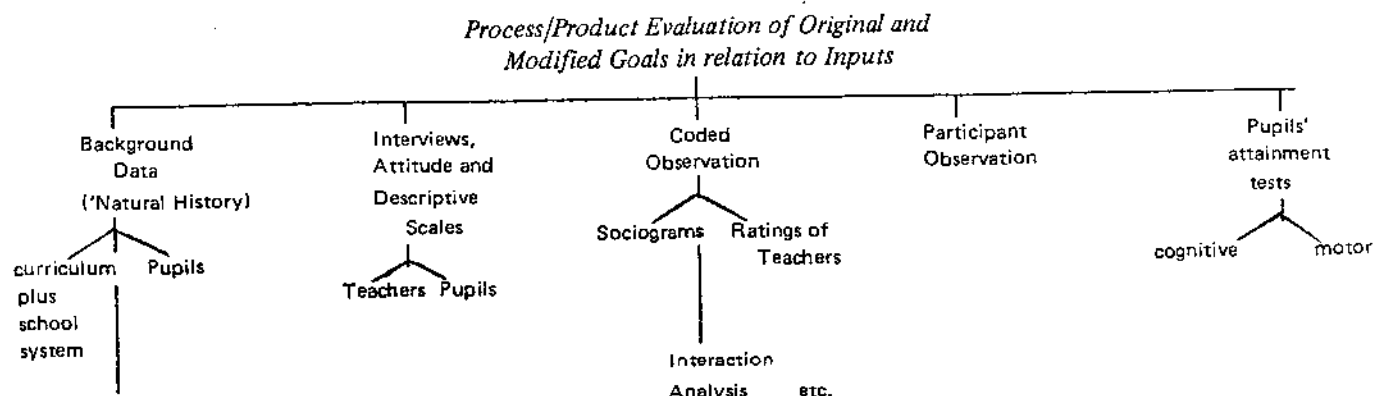
This reflects concern for the curriculum or instructional system in what the illuminators term 'the learning milieu'; it is, to use their jargon, both 'adaptable and electic', yet it stresses the necessity for multivariate measurement and analysis using the sophisticated statistical tools now at our disposal to test a multitude of hypotheses.

It would be naive of me to suggest the firmer incorporation of evaluations of teaching without recognising the possible human problems which this raises, particularly in England and Wales with out teacher-centred outlook. Evaluating teaching means evaluating teachers, and this is professionally for all of us in our social system, very threatening. However if the justification for schools and teachers ultimately rests in a concern for pupils and for society, then we may need to reconsider the nature and extent of some of the notional ethical boundaries which can insulate the majority of the teaching profession from extensive research enquiry. The science or art or craft of teaching, teacher training, and professional enlargement cannot possibly benefit from studies which exclude by design or political pressure investigations of the deep structure of teacher behaviour. I therefore conclude by suggesting that the evaluation of teaching has been a missing vital element in our evaluation plans; this has concealed some important factors affecting pupils' growth as a result of their curricular experiences. Is evaluating teaching a nettle which the curriculum evaluator must grasp?

Bibliography

A.E.R.A., *Monograph Series on Curriculum Evaluation*, Rand McNally, 1967 onwards.

Bannister, D. and Mair, J.M.N., *The Evaluation of Personal Constructs"*, Academic Press, 1968.

Comber, L.C. and Keeves, J.P. (eds.), *"Science Education in Nineteen Countries"*, Almqvist and Wicksell/Wiley, 1973.

Combs, A.W. et al., *"Florida Studies in the Helping Professions"*, University of Florida Social Science Monograph No.37, 1969.

Eggleston, J.F., in *"The Curriculum: Research, Innovation and Change"*, 99-113, Ward Lock, 1973.

Eggleston, J.F., Galton, M. and Jones, M., *Paedagogica Europaea,* **8**, 122-131, 1973.

MacDonald, B. and Parlett, M., *Cambridge Journal of Education,* **3**(2), 74-82, 1973.

Parlett, M. and Hamilton, D., *Occasional Paper No. 9, Centre for Research in the Educational Sciences,* University of Edinburgh, 1972.

Rosenshine, B., *"Teaching Behaviour and Student Achievement",* N.F.E.R., 1971.

Schools Council, *"Evaluation in Curriculum Development: 12 Case Studies",* MacMillan, 1973.

Walker, D.F., *Journal of Curriculum Studies,* **5**(1), 58-72, 1973.

SELECTING CRITERIA FOR EVALUATION

Lea Orr, NFER

(Paper presented to Symposium on 'Problems of
Evaluation'.

BERA Inaugural Conference, Birmingham,
April 6th, 1974)

I.   INTRODUCTION

Evaluation in educational research has been traditionally
identified mainly with curriculum. In this short paper I would
like to raise some questions connected with the evaluation of
a wider educational experience or 'activity', that is, the
evaluation of alternative educational institutions.

I shall do so with reference to a research project in which
I am presently involved. This is a study of the alternatives to
the traditional sixth-form, or rather of the educational provision
for the 16-19 age group. The project is now at its very first
stages, and therefore I am not in a position to report any findings
or conclusions. What I would like to do is to present the research
programme and the thinking behind it, to raise some questions of
methodology that have already been encountered and to indicate the
general approach adopted for solving or overcoming them. I hope
that these will stimulate comments and discussion which may be of
interest to this meeting as well as of benefit to the development
of the project.

II.   ALTERNATIVE INSTITUTIONS FOR THE 16-19

If the conventional grammar school sixth is thought of as
a model of the traditional sixth-form, alternatives have been
developing in recent years to many of its most characteristic
features. Its highly selective intake, its purely academic
and highly specialised courses, its strong emphasis on examina-
tions and university entry, its protective atmosphere, and its
generally elitist character; all these are being increasingly
challenged, at least in intention, by a number of trends in the
educational provision for the 16-19 age range. New types of
institutions are emerging within the schools sector and there
is a growing movement of 16 year olds from schools to colleges
of FE.

At present, within the maintained sector of the system there are five main types of institution offering full-time education to the 16-19 age group. Sixth-forms, in 'all through' grammar and comprehensive schools are the first two. These generally differ in their underlying philosophy and stated (or hidden) objectives, and in many cases also their range of courses and entry requirements, and hence in their intakes. A third type of institution is the sixth-form college which caters, under regulations, only for students aged 16-19 and may offer either purely academic courses, mainly at A-level, or a wider range of pre-vocational and non-examinations courses as well as O and A-level course. A fourth type of institution is the 'tertiary' college, which is a comprehensive college for the 16 plus, amalgamating all the sixth-form and further education provision in a given area into one college which is run under FE regulations. There are at present six such colleges in the country and a few more are definitely in planning. These offer the whole range of academic and vocational courses normally available in sixth-forms and in FE colleges and they are all 'open entry'. The last type is the college of further education (other than 'tertiary'). Apart from their traditional provision of technical and vocational education, FE colleges are increasingly providing full-time academic courses, including A-level, for increasing numbers of young students.

III. NEED FOR EVALUATION

The growing trend of movement to FE may be regarded as voluntary. The emergence of the new institutions, however, is closely associated with the reorganisation of secondary education along comprehensive lines. Essentially, they represent attempts, at the local authority level, to rationalise resources - of money, buildings and teachers - and to accommodate viable sixth-forms in the emerging comprehensive system while avoiding unmanageable school size.

These economic and administrative considerations figure, of course, in the public debate about institutional arrangements for the 16 plus. But they are strongly supported by educational, psychological and social arguments which concentrate on the advantages or disadvantages of alternative systems to the students themselves. It has been suggested that the wider choice of

ourses available in colleges, and the easy access of academic
tudents to related vocational courses, should benefit the
tudents educationally. Other characteristics of the colleges
uch as a break at 16, the less rigid selection at this age,
he more adult environment and the social mix of academic
nd vocationally based students should benefit them psycholo-
ically and socially. It is also argued that the existence of
he colleges is likely is likely to encourage staying on beyond
6. On the other hand, advocates of 11-18 schools claim that
he lack of continuity at 16 may discourage staying on, that
ollege students are deprived of the beneficial contact with
ounger pupils and also of better pastoral care, guidance and
ounselling, which are essential at this age.

These are only a few of the arguments brought forward.
he discussion is usually based, however, on statements of
faith with hardly any evidence being offered to support the
ropositions made. The most important feature of the current
system, which calls for an evaluative investigation of the
alternative institutions at such an early stage in their develop-
ment, is that many of the recent institutional and curricular
developments are seen by LEA's, schools and colleges as experi-
mental. The evaluation of their effectiveness and relevance
should help the existing institutions and also the many LEA's
which are still considering their plans for reorganisation.
Both planners and practising educationists are looking for
evidence and for lessons that may be learned from the experiences
of others. Such evidence is also likely to sharpen definitions
of objectives and so raise the level of argument among the many
interested bodies.

IV. FRAMEWORK FOR EVALUATION

The project has chosen the alternative institutional
arrangements as a framework within which the entire provision
for the 16-19 age group is to be analysed and evaluated. This
was done with full awareness that the criteria for evaluation
must be of content rather than structure. It is part of our
social experience, that structural innovations, when implemented,
may take different forms depending on the attitudes, expectations
and skill of the people who work in them. At the same time,
people's attitudes are effected by the structure of the institu-
tions in which they work. It is therefore clear from the outset

that we shall not find a one-to-one correspondence between types of institution and curricular or environmental variables. Nor do we expect clear cut relationships between the broadly defined structures and any measures of their output. However, as any curricular change requires some organisational change, new institutional arrangements may prove necessary for the implementation of curricular and environmental innovations. The alternative institutional arrangements were chosen as a frame of reference rather than, so to speak, as 'black boxes' whose inputs and outputs only are to be compared. We intend to look into these structures, to break them down into components, and to analyse the contribution of the components to outcomes. To put it differently, 'within group variance' with regard to the educational experience as well as to outcomes should be as central to the investigation as the variation between types of institutions. The model adopted for the evaluation exercise is that of assessing the inputs, outputs, and the educational 'treatment' with reference to institutional types.

This ought to include pupils' and teachers' attitudes and reactions to their institutions, as well as the image of these institutions in the eyes of potential customers, that is fifth formers who are about to decide whether, and where, to pursue their education.

A straightforward comparative study of the various institutional types is quite impracticable for other reasons. The deals of at least part of the different systems, and therefore their objectives, differ. Academic attainment is stressed by grammar schools and by many sixth-forms in comprehensive schools, while other comprehensive schools and many of the new colleges emphasise objectives such as a wider range of opportunities for students of a wider range of abilities and interests, vocational training, or social integration. The new colleges are few and evolving whereas many school sixth-forms are fully developed and may fall back on tradition. As far as the new colleges are concerned, it seems that their evaluation must take the form of a critical review of developing trends.

V.   OBJECTIVES OF THE SYSTEM AND CRITERIA FOR EVALUATION

Real difficulties start, of course, when criteria for evaluation are to be selected and defined. Special difficulties are bound to occur when post-compulsory education is under

consideration. To start with, there is an unbelievable lack
of consensus about objectives for this age range, not only at
the level of policy making, but also among practising educa-
tionists at the institutional level. There is no doubt,
however, that objectives, whether explicit or not, do exist
at any level, that they are multiple and in many cases are in
conflict with each other.

I would like to propose that in the context of the wider
education system, objectives for 16-19 education may be
classified under four broad headings:

(1) To contribute to personal development and social
     awareness.
(2) To provide a meaningful environment and to engage
     students in activities which are relevant and
     interesting to them.
(3) To prepare and select students for higher education.
(4) To prepare for employment.

An additional objective may be suggested, which unlike the others
does not apply to those already in the system and is probably
more controversial. That is:

(5) To encourage voluntary staying on.

Operational criteria for the evaluation of institutions can be
derived from the above classification

The 'staying on' objective may lead to the following
questions: who comes in; who does not, and for what reason; what
alternatives were considered; what range of courses is offered by
the institution; what are the entry requirements, both the formal
and the actual.

The 'meaningful experience' objective may be translated into
questions such as students' expectations at entry; their satis-
faction when on the course; the institutional arrangements for
change of course; facilities for social life; absenteeism; drop-
out.

Attainment of 'personal development' objective is naturally
the hardest to evaluate. However, questions on the following
lines may throw some light on this: access to varied experiences;

degree of specialisation; participation in the running of
students affairs, in the running of their institutions, in
social activities and in activities in the community; knowledge
of, and attitudes to possible educational and vocational careers;
institutional arrangements for independent study; general studies;
etc.

The two 'preparation' objectives, relating to higher educa-
tion and employment, lead to the criteria of attainment on
leaving the sixth-form; subsequent attainment when in HE or
employment; guidance and counselling; and satisfaction with the
sixth-form course in retrospect.

All these may be summarised under the following sets of
criteria: who comes into an institution and why; what does the
institution offer; how does it function; what is its underlying
philosophy and how is it being implemented; how popular is the
institution with its students, teachers and potential customers;
and last - attainment and satisfaction with the 'products'.

THE END(S) OF EVALUATION

David Hamilton

The Advent of Pluralism

Historically, the evaluation of educational programmed has been based upon assumptions, methods and criteria of relevance drawn from the experimental and mental testing traditions of psychology. Research within this framework has focused primarily on precise specification of objectives, behavioural measurement of achievement and statistical analysis of inputs and outputs. This research model derives from a 'systems' approach to education. Its governing assumption is that the goals of the system are self-evident or easily established. Concensus is taken for granted. If, however, concensus becomes problematic the model gradually begins to disintegrate. Eventually, it is replaced by alternative models that rely for their sustenance on notions of cultural pluralism. Models of this type accept the relevance of different audiences and even the antipathy of different interest groups. Now, as a consequence, we have open-ended, 'goal-free' evaluations that are 'responsive' to the 'trans-actional' relationship that exists between the evaluator, his clients and those evaluated. The emerging rhetoric quintessentially reflects the underlying constructs of the new order.

Yet, where does this lead the evaluator? By abandoning the security of objectivism for the ill-defined tenets of pluralism, have we opened the way to relativism? Shall we arrive at a position where there are no grounds for deciding the worth, truth or value of anything? Have we laid the basis for a value-free evaluation? If so, can we afford such a luxury?

The Competence of Evaluation

The second point follows on. Evaluation is an aspect of the change process. It relates, essentially, to decisions about courses of action that have been or are to be taken. Whatever the evaluator might claim, his work is seen as an integral part of that process. If there were no decisions, there would be no evaluation. Thus, evaluation becomes a crucial element in the mechanism used to procure change.

Inevitably, it is heavily implicated in the allocation of resources and, as a result, the politics of renewal. Different audiences and interest groups may have equal onto-logical status, but they are rarely equally powerful.

Concepts of cultural pluralism may be useful to an understanding of social phenomena, but are they an adequate basis for evaluation?

Pluralist Evaluation

The final point concerns the displacement of goals. Following the widespread acceptance of pluralist models for evaluation, its practitioners have progressively redefined their task. Generally, the trend has been towards diffuse-ness rather than specifity. In turn, the investigator has been constrained to adopt more holistic perspectives. In-creasingly, evaluators have become fascinated by the seamless cloak of education. Although such a dissipation of research energy may stem inevitably from the adoption of a pluralist stance, what comfort does it offer to the decision-makers.

To sum up: Educational evaluation may be at risk. There is evidence that it is developing into a subject without an object, a means without an end.

David Hamilton,
Glasgow University
  Department of Education,
4 University Gardens,
Glasgow G12 8QJ,
Scotland.

Discussion following "Evaluation" Session

After the four papers on evaluation had been presented, the discussion was thrown open to the sixty people who were present. Early contributions stressed the discrepancy of outlook between Dr. Hamilton's paper and others. Dr. Whitfield, in particular, queried the scenario that set "Illuminators" on the one hand against "Mechanists" on the other. He felt that the mechanists described by Dr. Hamilton certainly did not exist in the United Kingdom.

In reply, Dr. Hamilton cited examples of what he considered essentially trivial evaluative efforts, but Professor Wrigley felt that this was less than fair. Because of the widespread suspicion of the mechanistic approach, there had been difficulties in having evaluation included in British curricular reform projects. It had been included, but this was because those responsible had taken a broad view about what constituted evaluation.

Dr. Harlen remarked that curricular reform was now to be thought of principally as bringing short term changes in teachers rather than in pupils. She said that the psychometric tradition had initially misled us, and that insufficient attention had been paid to the teacher.

From the floor it was suggested that the whole methodology of curriculum reform was extremely shaky, and that resources would be better spent on the study of why particular children did not learn. Others, though, considered that some evaluative component in curricular reform was essential. The Twelve Case Studies recently published by the Schools Council were discussed and it was pointed out that every one of these studies originally started with the classical evaluation model, but later developed in other directions. Dr. Whitfield thought that much of this resulted from a shortage of cash. It would surely be better to collect all the relevant data at the appropriate time. He thought that the multivariate approach was very powerful, and had been insufficiently exploited.

This provoked an argument as to the number of variables that one would need in order to evaluate a program adequately. In general, people who used more variables did not seem more satisfied than those who used fewer. Dr. Pidgeon pointed out

that the 864 variables used in the recent IEA studies, to-
gether, accounted for only 40% of the between-students
variance in achievement.

At this point, a more basic division of opinion among
those present became apparent, and the use of multivariate
analysis was strongly criticised. The dependence of multi-
variate techniques on normative and parametric statistics
was seen as a deficiency. It was suggested that measurement
should be employed for exploring key hypothesis that might
"illuminate", and not for weaving multivariate blankets that
could effectively keep out the light.

The session closed with a plea for a more integrated
approach that would use the best features of various tech-
niques in a complementary manner.