DOCUMENT RESUME

ED 205 597   TM 810 502

AUTHOR   Hamilton, David
TITLE   Doing Justice in Evaluation Research.
PUB DATE   15 May 81
NOTE   22p.: Paper presented at the Annual Minnesota
              Evaluation Conference (2nd, Minneapolis, MN, May 15,
              1981).

EDRS PRICE   MF01/PC01 Plus Postage.
DESCRIPTORS   Case Studies: *Educational Research: Foreign
              Countries: *Program Evaluation: Research Design:
              *Research Problems: Social Values
IDENTIFIERS   Scotland

ABSTRACT
              The net result of the social-scientific developments
in the 19th and 20th centuries is that educational research has
inherited a science that is assumed to constitute a disinterested
technology of social engineering and a benevolent source of positive
social advance. Unfortunately, the social experiment conducted on
traditional lines depends upon the extent to which control can be
applied to the investigation: yet the achievement of such social
control, through the manipulation of "haves" and "have nots" may come
into conflict with new assumptions concerning social justice and
democratic freedom. The problem of truth or, to put it another way,
the problem of uncertainty is also ever-present in social research.
The debate over these issues will continue, but two ideas should be
kept in mind: (1) an evaluation is an intervention in a debate whose
terms of reference are never quite the same as those addressed by the
evaluation: and (2) the context of an evaluation or case study is not
the same as the context of the action that might flow from it.
(Author/BW)

DOING JUSTICE IN EVALUATION RESEARCH

David Hamilton
Department of Education
University of Glasgow
Scotland.

My role this morning is, I understand, to provide a prologue
to the ensuing, more specific papers. What I have to say, then,
relates to a number of topics. These include the purposes of
educational research; the role of social science in modern society;
and the significance of methodological debates in science. Overall,
my aim is to sketch in some of the context that gives meaning to
today's deliberations.[1]

My reasons for adopting this wide-ranging stance are two-fold.
First, it helps me to underline the fact thay many of the issues raised
about educational evaluation are also important to educational inquiry
in general; and secondly, it enables me to treat evaluation questions
not merely as matters of methodology but also as matters of ethics and
epistemology.

———————————

2

Innovation and Change

As a visitor to this country I never cease to be amazed by what appears to be a constant drive towards innovation in all spheres of life. Innovation is celebrated as a new frontier that, itself, is in a state of permanent renewal. Where I come from a different story obtains. Innovation is deemed a maverick, up-start activity. British innovators occupy the shady margins rather than the leading edges of society. Indeed, 'entrepreneur' is such an unspeakable title that the British had to turn to the French to find a label for it.

My contrast between conservative Britain and free-booting America is certainly overdrawn, though I think there is a real sense in which the innovation market and the turn-over of ideas is taken for granted in the USA. But there is another side to such an innovation-led economy. Does innovation necessarily promote change? Or is it the case that, under conditions of rapid innovation, change is more an illusory phenomenon - a social sleight of hand created through the sacrifice of substance to appearance, and fabric to fashion? Further, if this illusion is successful, are consumers lulled into a state of false consciousness? Do they assume themselves to be the beneficiaries of innovation? Or are they, in fact, its unwitting victims?

Let me bring my argument a little closer to education. Do the

changes we talk about really reflect some kind of directional shift towards a deeper or more immediate appreciation of educational phenomena? Or are they merely novelties that have been produced for sale on the academic market place? Is the topic of this conference, and our presence here today, the outcome of relatively random forces? are we just promiscuous followers of fashion, pursuing case study today, structural equations tomorrow, and something else the next day?

For my part I am sympathetic to such an analysis, but I do not believe that it provides the whole story. Day-to-day shifts in education may appear to be prompted by random or 'local' causes but, in the long run, a number of more profound if not irreversible factors also plays a part. That is, the dynamism of a nation's schooling is also affected by changes in a nation's economic structures, political systems, religious affiliations, cultural identities, and so on.

Doing science or doing justice?

Today I'm going to focus upon the relationship between educational inquiry and some of these broader shifts. The voices that I am goint to take seriously are those which, over the last twenty years or more, have argued the case for greater justice in social action. The burden of my argument is that the search for 'alternative approaches' in evaluation is part and parcel of the same movement.

4

Above all, I think the most important contribution made by recent
evaluation theory to the mainstream of educational inquiry has been
an increased awareness that 'doing science' is one thing, while
'doing justice' may be something else.[2]

The european scientific revolution of the seventeenth century,
with which we associate people like Francis Bacon and Isaac Newton,
was premised on the assumption that nature could be induced to give
up her truths if tackled in a new way; that is, through observation
and experiment. The scientific or empirical method as we know it
today can be traced back to that era. It constituted a new technology
that claimed to be more powerful and efficient than anything previously
envisaged. Mankind was no longer regarded as being in harmony with
nature. Henceforth, the relationship was to be one of domination and
subordination.

The extension of mankind's dominion over nature was the promissory
note issued by Bacon and his contemporaries. They believed that the
scientific revolution would usher in a period of great social wealth
and even greater personal and political freedom. In its most
democratically-organised form the scientific revolution promised a
heaven on earth for everyone. The social budget drawn up by its protagonists
assumed that technical efficiency would lead to the creation of a material
surplus which, in turn, could be redistributed to alleviate misery and

poverty for all time.

The Social-Scientific Paradox

The aspiration to control nature, however, also triggered a related sentiment - that a comparable technology could be developed to control the workings of civil society. The nurturance of this idea was a key feature of the 18th century Enlightenment; and its rise to maturity in the 19th century was intimately tied up with the emergence and establishment of such disciplines as economics, sociology, psychology and, not least, education.

With the wisdom afforded by hind-sight, I think we now recognise that the social-scientific revolution was built upon a paradoxical core assumption - that the purpose of the social sciences was to extend mankind's domination over mankind. Before the nineteenth century such an assumption was unproblematic: a large proportion of society was denied the civil rights enjoyed by the remainder. Thus, the original purpose of the social sciences was perfectly consonant with a 'two nations' image of society. Accordingly, social scientists were able to develop a rationale and technology which, explicitly or implicitly, took only indirect account of the feelings and aspirations of those whose lives they intended to reform.

It was a time when the expert 'knew best' and a time when the administrative solutions of the experts were deemed to be 'in the

6

best interests' of their clients. The justice of such methods
was underwritten by an appeal to the principle that what is good
for society is also good for every individual; and the plausibility
of such methods was assured by the fact that they operated within
expanding national economies that were always able to provide a
little extra for everyone.[3]

Scientific Method as Problematic

The net result of these nineteenth and twentieth century
movements is that educational research has inherited a science that is
assumed to constitute a disinterested technology of social engineering
and a benevolent source of positive social advance.

I am sure that it will be no great surprise to you that I am
sceptical of this ahistorical view of science. In particular, I
discount its claims to have found a set of universal or absolute
standards - 'the' scientific method - against which other forms of
inquiry are judged and, usually, found wanting. My own view
is that any methodology like any course of action, necessarily
has an ethical dimension and that, accordingly, if we change our
ethical criteria we are inexorably led to a reconsideration of our
practice.

Thus doing evaluation is not the application of a universal
method but, rather, a constraint striving to resolve - for every new
occasion - the essential tension that suffuses our concept of social
science. And it is this striving that is one of the subtexts of
today's conference.

Let me amplify my argument by taking a closer look at the
'universal or absolute standards' that I referred to a moment
ago. In particular, I would like to demonstrate certain problematic
features of our current use of the concepts of efficiency, truth
and objectivity in research.

Efficiency and the 'True' Experiment

Nearly twenty years ago Donald Campbell and Julian Stanley
focused on the problem of efficiency in their seminal paper on
'Experimental and quasi-experimental designs for research on
teaching'[4]. Their work, quite explicitly, was an attempt to
extend the rationale of biological experimentation into areas where
experimental control is less than complete. Two features ofCampbell
& Stanley's paper are pertinent to my argument. First, it was a
fomal analysis, having more regard to the design of experiments than
to their implementation; and secondly, through its discussion of
external threats to validity[5], it noted thatthe results of 'true'
experiments may be open to multiple interpretation. Put another
way, their paper was about the reduction of uncertainty in the
social sciences, not its elimination.

A few years later Campbell revisited these earlier ideas in
a paper on 'Reforms as experiments'.[6] His particular interest was
the contribution that the social sciences might make to the 'Great
Society' legislation of the 1960s. His paper concerned itself with
the 'political realities'[7] as well as the technical logic of
experimentation. And, by the same token, it was addressed directly

8

to the reforming administrators (or 'social engineers')
of that epoch.

In the course of his argument Campbell ran up against a general
problem in experimental research : the social experiment conducted
on traditional lines depends upon the extent to which control or
randomisation can be applied to the context of the investigation;
yet, the achievement of such social control, through the manipulation
of 'haves' (the treatment group) and 'have nots' (the control group)
may come into conflict with new(or Great Society) assumptions
concerning social justice and democratic freedom.

In the late 1960s, then, Campbell noted the intractable nature
of the problem but, at the same time, came to the defence of the
experimental model when he claimed to have identified a set of
experimental conditions wherein randomisation was not a variance
with social justice. That is, he had found a situation where, even
in Great Society terms, the efficient experiment is the same thing
as the just experiment. The particular design requirements noted
by Campbell relate to the allocation of 'social ameliorations' that
are in 'short supply'.[9] Campbell argued that under such conditions
it would be impossible for everyone to receive the benefits of the
treatment and that, therefore, randomisation was both technically
and ethically defensible .

Campbell's a. .aent - based on a discussion of the Salk vaccine trials-
is very plausible, but it is flawed on at least one count: the Salk
trials are not a general instance. On that occasion, the experimental

subjects received a treatment of presumed benefit, whereas the control subjects were deprived of something - immunity to polio - which in fact, they had never enjoyed. Thus, contrary to his claim, Campbell's exemplar only embraces instances where benefits are randomly with-held, not instances where benefits are randomly withdrawn. In the second case, of course, the deprivation is much more visible to the deprived.

For these reasons, then, I would expect randomisation experiments involving increases in grant aid to provoke a different consumer response from those than entail cuts in grant aid. In short, the research designs proposed by Campbell may be logically equivalent, but they are not socially equivalent. Despite his claim that 'this is the ideology that makes possible "true experiments" in a large class of social reforms',[10] Campbell's attempt to reduce social justice to social efficiency must be regarded as unsuccessful.

The net result of Campbell's efforts, of course, is that a twentieth century tension is still resolved in favour of a nineteenth century conception of the relationship between science and justice. That is, 'organisational efficiency' and the 'scientific organisation of society'[11] are still deemed to be at peace with the interests of

10

the individual.

In later work Campbell has drawn back from some of his
earlier positions and has been less ready to settle in favour
of managerial models of control and efficiency. For instance,
in a paper directly related to the themes of this conference -
'"Degrees of freedom" and the case study'[12] - he has discussed
the value of adopting procedures which, unlike those of 'true'
experimentation, deliberately allow 'innumerable alternative
solutions'[13].

But Campbell's new-found support for case study methods still
relied unilaterally on technical criteria. His argument was that the
convergent nature of experiments means that only a relatively limited
number of hypotheses can be examined whereas, by its multi-dimensionality,
a case study allows many more explanatory hypotheses to be
scrutinised, including those that were not envisaged at the outset
of the research.

Truth in Research

But, you might ask doesn't the possibility of alternative solutions
also include the possibility that no one solution will emerge as
'the' explanation. And doesn't this possibility in its turn, undermine
the expert status of the researcher? This question brings me to the second
issue I raised earlier - that of truth in research.

The problem of truth or, to put it another way, the problem of uncertainty is ever-present in social research. It even exists, as I showed earlier, in the organisation and interpretation of "true" experiments. In an important sense, however, the uncertainty does not reside in the phenomena that we are trying to study but in the way we value them.

Let me offer an example. Imagine two young children: one that can count up to five and the other that can count up to ten. Six weeks later the first one has moved up to seven digits and the older one has moved up to thirteen. The question that might be asked of these children is 'which one has made the most progress?'. At one level this question is absurdly simple; but at another the answer given can only be arbitrary. As this example suggests, the facts do not speak for themselves, they merely beg further questions.[14] In the instance cited, any answer will hinge upon whether the initial differences between the children should be taken into account. That is, it depends on the moral valuations that we bring to the task. Do we treat the children and their circumstances as different or equal? How, in technical terms should we 'weight' them?[15]

The traditional reaction of educational researchers to the challenge of these and other empirical uncertainties has been to treat them as problems of technical rather than ethical clarification. It is assumed that they can be solved by using a higher level of

methodological sophistication (e.g. by using multivariate rather than bivariate statistics). The net result, of course, has been the pursuit of a spurious precision in research; itself driven by the popular but mistaken belief that the problems of research can be reduced to the realm of methodology and technique.

Such activities are easily driven forward. Their technical precision susstains an illusion of certainty. But, as in life, there is always the risk that procedural consensus will break down, pulled apart by the different valuations adopted by the protagonists. At first glance, this possibility seems remote. Yet it is the very issue that has fueled the debates and recriminations surrounding the Head Start and Follow Trhough evaluations.

Whenever we hear the call for results do we have sufficient resources to back up Bacon's promissory note? Are our procedures as 'hard' as we would like to imagine? Does our delivery record match up to our funding talk? I think not. In practice we operate with standards that are more double than absolute. Our professional misgivings about the certainty of our procedures are, more often than not, expressed as public silences. Our audiences, therefore, still expect us to furnish a kind of 'out there', theory-free truth which, increasingly, we know we cannot provide. Must we continue in our cover-up? Or can we offer an alternative view of the social sciences?

Objectivity and the Independent Observer

The professional misgivings and public silences that relate to
conventional notions of truth also apply to objectivity - the last
issue that I would like to address this morning. If education, like
physics, is regarded as an exact science then objectivity and exactitiude
are taken to be synonymous. Researchers are assumed to occupy a
privileged vantage-point - one that enables them to render accounts
of the world that are uniquely free of error or bias. Hence, in
these terms, a drift towards subjectivity is, at the same time, a drift
away from the truth. But physics provides us with an alternative
perspective on truth and objectivity. Since the publication of
Einstein's Special Theory of Relativity, in 1905 it has been more
readily appreciated that any description of the world is dependent
upon the time/space location of the observer. Thus, the objectivity
of an observation, like its truth, is always dependent upon the personal
knowledge or, what amounts to the same thing, the ethical standpoint
of the observer.

These assertions of a physicist about the convergence of truth
and subjectivity may seem a little remote from the hurly-burly
of education inquiry. But, again, they served as an important (if
uncited) precurser of Donald Campbell's paper on 'Qualitative knowing in
action research'[16] that was read to the American Psychological
Association in 1974. The purpose of Campbell's address was to revisit

some of the issues that had been left out of his earlier discussions
of 'quantitative-experimental' approaches to social research. What
is even more important to this conference is that he included
'case study'[17] as contributing towards the 'qualitative knowing'
of the social sciences.

Set against Campbell's earlier papers on 'Experimental and
quasi-experimental design' and 'Reforms as experiments',
'Qualitative knowing in action research' has two noteworthy features.
First, it disavows managerial models of control by counterposing
the value of 'participant evaluation'; and secondly, it takes the
convergence of subjectivity and truth to such a point as to
suggest that the experiences of 'participants' are more 'valid'
than those of 'outside observers'[18]. Overall, Campbell's
'qualitative knowing' paper is significant in that it demonstrated,
for him at least, a hitherto unacknowledged awareness of the
essential tension that pervades the social sciences. Whereas Campbell's
initial methodological strictures arose from an apparent concern
with efficiency, by 1974 the concern for social justice was more
obtrusive. But, to this day, the problem of combining a 'top-down'
model of efficiency with a 'bottom-up' model of justice seems
to have proved as intractable to Campbell as the earlier problem of
divining the 'true' experiment.

Towards an Alternative Social Science

Let me pause for breath. So far in this paper I have identified

a tension that, I believe, is inherent to social inquiry in general and educational inquiry in particular. I've noted too that working in the context of expanding economic systems social scientists have been successful in drawing a technical veil over these issues.

Today, I have tried to pull back a little of the veil. I have questioned some of the assumptions that have guided social science practice for many years. Together they underwrote a science - a technical-ethical calculus - that, somehow, was deemed to be beyond society. Hence when the science failed, it was the earth-bound practitioners who were blamed. 'You didn't do it right' is the charge that was made against them But it is not our world that fails to live up to our science- rather, it is our science that fails to live up to our world.

Where, then, does that leave us? I suspect that if you came to this conference worried about methodology you are feeling even less comfortable by now. You came looking for solutions and all I have been able to offer are irresolvable ambiguities and infinite regressions. First you saw black boxes in research ; now they have turned into black holes. But, take comfort, you are not being deluded.

In a very real sense, methodological debates are no such thing,

That is, they are not purely about technical matters. Instead,
they are driven by all kinds of theoretical, ethical and social
concerns. Indeed, the lack of consensus that create; such a
debate is, I believe, a reasonable response to the many different
ways in which the efficiency/justice tension can be resolved in
the social sciences. But where do these ideas spill over into
the realm of 'case study methodology in educational evaluation'?
Let me to conclude, suggest two points of departure.

First, I find it useful to consider an evaluation - formal
or informal - as an intervention in a debate whose terms of
reference and criteria of worth are never quite the same as those
addressed by the evaluation. This may arise because the evaluation
only addresses a subset of issues or because the debate, itself,
is constantly changing. Either way, the evaluation can never replace
the debate; that is, it can only be a sensitizing device. It
always remains 'an' evaluation, not 'the' evaluation.

And second, I find it useful to remember that the context of
an evaluation or case study is not the same as the context of the
action that might flow from it. That is, other circumstances and
other audiences may intervene. Necessarily, then, a case study
evaluation can only prefiture 'a' course of action, not 'the' course
of action.[19] If we try to offer any more than these gentle promises
we too, may be guilty of a double standard in our work.

In preparing this paper, I worried whether I was devoting enough

space to case study and evaluation. Now I suspect I've said

too much. Nonetheless, I hope my remarks provide a suitable back-

cloth to the remaining papers. My belief is that there are

substantial links between the general topics of my paper and

the current concerns of evaluation research and case study

inquiry. I hope, therefore, that what I have said contributes

towards building a bridge between your own experiences and those

that are to be examined in the presentations that follow.

1. For two volumes that present more exhaustive analyses of case study and evaluation see H. Simons (ed) <u>Towards a science of the singular</u>. Norwich, England: Centre for Applied Research in Education, University of East Anglia, 1980; and T.S. Popkewitz & B.R. Tabarchnick (eds) <u>The study of schooling: field based methodologies in educational research and evaluation</u>. New York: Prager 1981.

2. For an extensive discussion of justice in research see E. House, <u>Evaluating with validity</u>. Beverly Hills: Sage, 1980.

3. It is no accident that our social science paradigms took shape in France, Germany, Britain and the USA. It was only in such economically-powerful countries that the essential paradox of the social sciences could appear to have been resolved. For a more extensive discussion of nineteenth-century social science see D. Hamilton, Educational research and the shadow of John Stuart Mill. In J.V. Smith & D. Hamilton (eds), <u>The meritocratic intellect: studies in the history of educational research</u>. Aberdeen: Aberdeen University Press, 1980, 3-14.

4. See N. Gage, (ed), <u>Handbook of research on teaching</u>. Chicago: Rand McNally, 1963, 171-246.

5. The distinction between internal and external validity seems to
    date from D.T. Campbell, Factors relevant to the validity of
    experiments in social settings. Psychological Bulletin, 1957,
    54, 297-312.

6. The original version of 'Reforms as experiments' appeared in the
    American Psychologist, 1969, 24, 409-429. This paper uses the
    version in E.L. Struening & M. Guttentag. (eds) Handbook of evaluation
    research (vol.1). Beverly Hills: Sage, 1975, 71-100.

7. Ibid., p.72.

8. Ibid., p.84.

9. Ibid., p.86

10. Ibid., p.87

11. Ibid., p. 72

12. Comparative political studies, 1975,8, 178-193.

13. Ibid., p.184.

14. For a witty and penetrating critique of the notion that the facts speak for themselves see K. Harris, Education and knowledge. Boston: ι Routledge and Kegan Paul, 1979 (especially pp. 10-18). Also see H. Freeman & A. Jones, Educational research and two traditions of epistemology. Educational philosophy and theory, 1980, 12, 1-20.

15. For an early and lucid discussion of the practical and ethical issues in the assessment of learning see R. Stake, Measuring what learners learn. In E.R. House (ed) School evaluation, Berkeley: McCutchan, 1973, 193-223.

16. In this paper I have used a mimeo version of the original talk.

17. Ibid., p. 1.

18. Ibid., p. 26.

19. For a more extensive discussion of evaluation that is in tune with most of the ideas in this paper see S. Kemmis, Seven principles for program evaluation in curriculum development. School of Education, Deakin University, Victoria, Australia, 1980 (mimeo). Kemmis builds his argument around the thesis that 'evaluation is the process of marshalling information and arguments which enable interested individuals and groups to participate in the critical

debate about a specific program'. See also Ruth Anne

Olson's work in Minneapolis described in _Evaluation as_

interaction in support of change, Grand Forks, North Dakota;

North Dakota Study Group on Evaluation, 1980.