THE VERIFICATION OF DESCRIPTIVE CASE STUDIES.

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

Classic experimental designs used in educational research in field settings, which involve population sampling, experimental control and systematic correlation, can in principle test theories and, in such activities as agriculture, can guide action. The guidance they offer to action is, however, limited by the fact that it is actuarial, dealing in probabilities, offsetting gains and losses; and the assumption is made that consistency of treatment will be maintained throughout a population. Thus the same seed strain and fertilizer mix will be used throughout a field.

There are difficulties in applying this research paradigm in education. First, there are limits to the principle that the educator can offset gains and losses, following a probabilistic generalization rather than his professional judgement, and discount the damage done to some students by the benefits to others. And, second, it is possible in education to differentiate treatments by school within a school system, by class within a school, by pupil within a class. Educational action is acutely sensitive to ecological factors, to historical setting and to individual variability, and consequently the same procedures cannot be recommended across cases. Judgement is needed to differentiate action according to circumstances. Hence research must be elucidatory not merely evaluative. It must feed judgement in the diagnosis of the case and the prescription of the treatment.

Although it might be possible in principle to feed such judgement by reporting theoretical findings based upon the results of the
statistical analysis of trait-treatment interactions, the complex of variables seems so intricate that it eludes the techniques available to us.

It was experience of the difficulty of this enterprise which led Lee Cronbach, in an address to the American Psychological Association in 1974, to try to reach 'Beyond the two disciplines of scientific psychology'. Expressing scepticism about the possibility of a nomothetic science of education, he suggested the need for interpretation in context as contrasted with generalisation. He had this to say:

The two scientific disciplines, experimental control and systematic correlation, answer formal questions in advance. Intensive local observation goes beyond discipline to an open-eyed, open-minded appreciation of the surprises nature deposits in the investigative net. This kind of interpretation is historical more than scientific. I suspect that if the psychologist were to read more widely in history, ethnology, and the centuries of humanistic writings on man and society, he would be better prepared for this part of his work. (Cronbach 1975, 125).

The psychologist who did turn to humanistic writings to ethnology and to history would, if he were perceptive, find not escape from discipline, but hints towards alternative disciplines, and it is these hints I shall follow up here. In my view they have not been given the attention they deserve by the majority of those who, over the past few years, have reported or discussed naturalistic research or school ethnography.

In this research tradition the researcher is involved in fieldwork, in which the principal components are participant observation, non-participant observation, interview and the collection of documents. At the conclusion of the work in the field the researcher commonly writes a report which is a 'portrayal' (Stake 1972), a reasoned
summary of conclusions or a theoretical interpretation.

The question arises how the reader-fellow researcher or parasite doctoral examiner - can test critically the reported research; that is, how to handle the problem of verification. This problem is one to which I gave attention in a paper read at the American Educational Research Association meeting in 1977 (Stenhouse 1978) and later in a paper given to the Scottish Educational Research Association (Stenhouse 1979). Currently, I am directing a project on Educational Case Records supported by the British Social Science Research Council, which is largely concerned with this issue.

In the present paper I shall look briefly at the problem of verification in psycho-statistical experiments and then review the same problem in humanistic fictions, in ethnography and in history. Finally, after a brief consideration of the main issues, I shall propose conventions to support the verification of work in naturalistic case study.

Verification in Psycho-statistical Experiments

I shall define my use of the term, 'psycho-statistical' experiments by reference to Fienberg (1977) from whom I adopted the adjective, and to Fisher (1935) and Campbell and Stanley (1963). The tradition so defined is, I shall assume, familiar enough.

There are two points at which we can ask questions about verification: first, does the result of the experiment seem to be trustworthy? second, does the result of the experiment make sense in terms of theory? The first of these may be regarded as a principle of replication and the second as a principle of coherence.

Although in principle replication depends upon the same result being recorded if the experiment is repeated, in practice comparatively few field experiments in educational research are in
fact replicated. Rather the design and conduct of the experiment is criticised in context and an informed public makes judgements as to whether replication would yield a similar result.

As Gene Glass put it in a discussion with the author, the position taken is: 'Don't ask of my results do they persuade you? are they coherent? is the picture clear? do they appeal to your common-sense? Just examine my methods. If my methods were correct, I can't falter there' (Glass 1977 interview with author: Glass is characterizing, not defending, this position).

Underlying this position is a principle I shall call estimate of replication: that is, providing in terms of an assessment of the design and conduct of the research an answer to the question: 'Would another researcher (perhaps, would I) get the same result if he designed and executed his experiment as nearly as possibly perfectly?' This is, of course, an extension of the notion of reliability in the face of a particular formulation of the problem.

In spite of Gene Glass's mention of the disparagement of coherence in some educational research circles, coherence is an important principle for verification; but in the experimental paradigm it depends upon the embodiment of coherence in articulate theory, structured with sufficient clarity for hypotheses predictive of experimental results to be deduced. It is this condition which makes possible the elegance and parsimony of the crucial experiment. However, theory of this kind, which is rather highly developed in the hard sciences, is none too easy to come by in the social sciences and is particularly elusive in education. More often, we are reduced, as Glass hints, to commonsense.

The role of coherence in verification is central to classic scientific method: if results and theory are in conflict, then theory must be revised or results must be discredited. Coherence and replication go hand in hand. And coherence is an extension of the notion of validity.
In short the highest credence will be given to a result if the experiment is well designed and the outcome is consistent with established theory and/or commonsense. Less credence will be given to the result, the more flawed the experimental design and the more paradoxical the finding in terms of established theory or commonsense.

It is also important to point out that this logical position is embodied in the academic community. These are not merely abstract principles, but principles expressed in the deliberations of scientific conferences such as this. And no academic community can exist without a division of labour which means giving some credence to the judgement of others, which recognizing that judgement, and not objectivity, is involved.

A considerable problem is posed when the intelligence, personal confidence and technical expertise of individuals becomes daunting to critics. The case of Cyril Burt reminds us to distrust the results of the researcher who frightens off critics.

**Verification in Humanistic Writings on Man and Society**

By 'humanistic writings' I take it we mean novels, drama, poetry and, by extension beyond the written, cinema and television. Other philosophical and reflective writings might be considered, but they are distant from the descriptive portrayals of social life with which we are concerned in case study.

The humanistic writings which are like case-study portrayals are distinguished from such portrayals by their status as fictions, but they are in some sense realistic fictions. Here the term fiction signals not that they are false, but that they are products of imagination. And such products of imagination have tests for truth: that is, they involve a problem of verification.
The full explication of such a view - that fictions of this sort present truths - is beyond the scope of this paper. And here we can restrict ourselves to the realistic truth in fictional portrayals, though that is by no means the whole story. This is the truth we imply when we praise Shakespeare or Flaubert for their insight into people.

Natural perceptions of human experience and interpretations of behaviour are always perceptions thrown into a perspective by a point of view which is defined by the social and personal attributes of the observer. It is the essence of the arts that they offer forms in which such perceptions can be played with by the artist. The conventions of narrative portrayal, for example, not only allow the writer to invent his story and his narrator, but, if he wishes, to adopt the convention that he can read minds, and to take other liberties.

Since the writer has these liberties of fiction at his disposal, he is not required to authenticate his portrayal as factual and grounded in specific empirical evidence. Indeed, it is part of the function of fiction to allow a portrayal which may offer both imagined perceptions of a world outside the perceiver and imagined subjective experiences of a kind to which in principle no author and no reader can have access. And though the author has a 'method', this method is expressed as 'form', which, when it is successfully employed, is not properly divisible from a content of results. Consequently, a critique of method is not a basis for an estimate of replication. The very terms we are using begin to collapse under us.
In realistic fictions verification depends upon the readers' judgements about the trustworthiness of a portrayal in terms of verisimilitude: a likeness to factual truth. Such judgements are estimates of coherence between the portrayal and the experience of the reader. Some would characterize this experience as tacit theory, and I believe it has that aspect since we clearly generalize beyond our first-hand experience into imagining likelihoods. But another aspect of judgement of verisimilitude is the capacity of the brain to conduct a rapid search of memory evoked by a fictional stimulus which provides an image which, by virtue of this power, we call 'evocative'.

The verification of a fiction in terms of verisimilitude depends upon appeal to what I shall call a 'general second record', that is, the organized trace of the reader's experience. Coherence between the fiction and the general second record confirms the fiction. Conflict places the portrayal under criticism, and produces a situation akin to that in science when an experimental result challenges established theory. Either the result must be proved untrustworthy or the theory must be revised. In a fictional humanistic account which is not confirmed by the general second record either the portrayal must be dismissed or it must be upheld by a revision of the record of one's own experience. And people who discuss the verisimilitude of fictions actually make an appeal to each other's experience in the belief that some common grounds of human experience can be found.

There is, in short, an established intersubjectivity of criteria by which realistic fictional portrayals are criticised in terms of verisimilitude. Like theory in the social sciences, this intersubjectivity is not unitary or fully integrated, but it enters into all judgements about portrayals. I have given the personal subjective representation of this intersubjectivity the rather odd
name of a 'general second record', borrowing the term 'second record' from J.H. Hexter (1972) because it will help to relate this treatment of the verification of fictions to the account of verification in history which follows.

Verification in History

Like fiction, history often tells a story, and the historian, like the novelist, interprets life. History is sharply distinguished from fiction, however, in that interpretation is mainly a matter of selection and commentary, lacking the resource of invention to the extent that the factual and experiential bases of history must be authenticated.

Although artefacts can be a basis for authentication in history and are the only basis for the authentication of prehistory, history is traditionally a documentary science: it rests upon written records which are either the surviving by-products of social process - parliamentary transactions, company balance sheets, minutes of meetings - or the witness of individuals specifically written for audiences - the diary, the account of the voyage, the apologia of the condemned man.

Historians attach particular weight to the 'primary source' - that is the document providing first-hand evidence of the past or the material evidence of buildings and the kinds of objects one finds in museums. This is the fundamental raw material of history, which it is the task of the historian to interpret. The primary sources are sometimes supplemented by secondary sources - second-hand accounts of the past - but these are less acceptable as foundations for an historian's account though extensively used in that history which is pitched at so general or broad a level that
its author must necessarily build on the work of other historians. Together the primary and secondary sources constitute the first record which is contrasted with the 'second record' already alluded to above.

The first record provides the given or data of history. I prefer the word evidence to data here, because it seems useful to distinguish data, as being gathered by methods of observation or collection which keep in mind the problem of comparability or standardization, from evidence, which is problematic as to comparability and requires critical appraisal and comparison to establish its status and the relationships within it. There is in history nothing like the control of measurement and experiment which reduces the scientists evidence to data.

In history - as in science - all 'facts' are inferences, and though some inferences are fairly direct and simple, many are much more contentious. And the facts which are inferred are selected and interpreted. It is in the processes of inference, selection and interpretation that the historian draws upon his second record.

Baxter (1972), who coined the term, writes of the historian's second record as

everything he can bring to bear on the record of the past in order to elicit from that record the best account he can render of what he believes actually happened in the past. Potentially, therefore, it embraces his skills, the range of his knowledge, the set of his mind, the substance, quality and character of his experience - his total consciousness. Since no historian is identical with any other historian, what each historian brings, his second record, differs in some measure from the second record of every other historian. (p.104).
Verification in history depends upon a kind of replication in respect of the first record which in turn rests on a coherence of intersubjectivity in the second record.

The first records or primary sources of history are essentially documents or artefacts in the public domain and the inference of facts from these sources is verified by virtue of their being accessible to a number of professional historians who discuss the status and implications of the first record. Thus replication is a matter of revisitation of the record. We believe that a particular historical event actually happened because we know that many historians have investigated the matter from primary sources and have mutually confirmed each other. There is some analogy here with the faith in the results of the well-designed experiment which we discussed earlier, and the signposts of the well-designed historical investigation are the footnotes which lead back to the first record. This first record is the foundation on which history is built.

Consensus between historians in its interpretation - and, indeed, even the capacity to dispute interpretations by discussion - depends upon some intersubjectivity in the second record. In spite of Buxton's correct assertion that the second record of each historian is in some respects unique to him, there is also much common ground.

Some of this common ground lies in what I earlier called the 'general second record' of experience of life in a given culture. But there is also a professional second record, accumulated by the historian through the experience of studying history rather than through the general experience of living. When the historian says: 'People just don't behave like that', he is drawing on his general second record; when he says: 'Fifteenth century churchman just don't behave like that', he is drawing on his professional second record.
Verification in history, as I have described it, may be expressed in tabular form as in the figure below, but it must be borne in mind that all assessment of evidence for trustworthiness depends upon and interacts with interpretation. The four cells of the table represent simultaneous interactive operations and not alternatives.

Verification in History

<table>
<thead>
<tr>
<th>Evidence assessed for 'trustworthiness'</th>
<th>Interpretation assessed for 'coherence'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Sources</td>
<td>Professional Second Record</td>
</tr>
<tr>
<td>Secondary Sources</td>
<td>General Second Record</td>
</tr>
</tbody>
</table>

Verification in Ethnography and Ethnology

Classic ethnography was forged in the study of preliterate cultures, the systems of meanings which sustain the conduct of living in societies which, from the point of view of the western social anthropologist or ethnologist, are 'ethnic' and unfamiliar. The ethnographer, who studies such a society, is a stranger in it, and the technique he uses for his study is 'participant observation'.

There is an extensive literature on the participant observer's role as a research tool, but its basis can be relatively simply
explained. The participant observer takes part in the life of the community he studies, learning the language and customs in the process, but at the same time retaining his own perspective from which to study the life of that community rather than simply adopting the culture and 'going native'. His observations of life and behaviour in the community, his conversations or interviews and his observation of his own experience of the subjective correlates of the life of the community are carefully recorded in 'observer's notebooks'. Such notebooks are - or ought to be - written up regularly as soon as possible after the experiences they record.

Now one might argue that, if we use the terms we have introduced above, the observed life of the community is the primary source and first record of the ethnographer and that the observer's notebook is a secondary source produced by the operation of the second record on the first record. Then by analogy with history verification might be a matter of revisitation of the first record by different scholars. Indeed Claude Levi-Strauss looks at this view.

... history and ethnography have often been contrasted on the grounds that the former rests on the critical study of documents by numerous observers, which can therefore be compared and cross-checked, whereas the latter is reduced, by definition, to the observations of a single individual.

To this criticism we reply that the best way to overcome this obstacle in ethnography is to increase the number of ethnographers.... Today there are indeed few peoples who have not been studied by numerous investigators and observed from different points of view over a period of several generations, sometimes even several centuries. (Levi-Strauss 1977, 17).
Bearing in mind the element of change in preliterate societies, I am doubtful if this approach to verification in ethnography can be carried far.

Levi-Strauss, though he writes of the anthropologist's developing interest in applying his methods and approaches in the study of advanced societies, notes that

the absence of written documents in most so-called primitive societies forced the anthropologist to develop methods and techniques appropriate to the study of activities which remain, for that reason, imperfectly conscious on all the levels where they are expressed. (Ibid, 24)

Because he takes the lack of writing or of direct comment to define unconsciousness of the unrecorded elements in social life, Levi-Strauss sees anthropology as concerned primarily with the unconscious elements in social life, while history is concerned with the conscious. This is not a hard and fast distinction, but it does have a heuristic power, and it leads Levi-Strauss to the declaration that

The anthropologist is above all interested in unwritten data, not so much because the peoples he studies are incapable of writing, but because that with which he is principally concerned differs from everything men ordinarily think of recording on stone and paper. (Ibid 25)

But the ethnographer does accumulate written records - his own notebooks - and would be thought vulnerable to criticism if he did not, and he writes up his work by a retrospective analysis which reviews his notes in the light of a second record. If they are
not a simple and direct primary source, the ethnographer's field notes are the nearest ethnography can offer. And the interpretation of field notes is clearly crucial for ethnography. Yet ethnographer's field notes are not published or placed in the public domain.

The foundations of ethnography were laid by traditional European gentlemen in the nineteenth century. As has been wryly remarked: 'It's a good job they were gentlemen because we've only their word for it!'

This is not quite fair comment, but there is a grain of truth in it. If we take our table of verification in history as our starting point, then we might say that the ethnographer's field record is the primary source and the ethnographic account of a culture in its reported form is a secondary source. And certainly for the most part it is only the secondary sources which the ethnographer makes available to other ethnographers and ethnologists. To some extent this is understandable because the economics of publishing have not made it feasible to print ethnographers' field note books. However, the recent advent of cheap microfiche publishing raises the question: 'Should not ethnographers' field note books be made available more widely in this form. Just how widely is a matter for further debate.

Rather more complicated is the problem of the second record of ethnology. For certainly the general second record, which is fed by experience of living in our culture, is called into question by ethnography. To the protest: 'People just don't behave that way', may come the response: 'Not in your culture, but in this culture they do'.
In fact the case is not quite so sharp as this makes it sound, for, while it is difficult to dismiss an ethnographic interpretation by reference to the second record, it is to some extent inevitable that the general second record confirms the ethnographic account. For instance, the Kwakiutl potlatch, in which members demonstrate their wealth and status by burning their possessions, rings true as a parody of conduct in our own society. It remains true, however, that the general second record is weak in a study which turns its attention professedly to the study of the exotic. There is an underlying character of 'Believe it or not' about ethnography.

The defects in the general second record are to an extent made up by an extension of the ethnologist's professional second record. However, this second record is not like that of the historian who 'knows his period' in company with his fellow historians. The ethnologist whose fieldwork has been conducted among the Bantu must apply his second record to a critique of an ethnography of the Lapita; and he does. This application of the professional second record is far more theoretical and technical than that of the historian. It is founded on the idea that certain features generalize across cultures, and hence that certain concepts, such as kinship, ritual, rites de passage, can consequently be applied across cultures. These concepts - unlike 'parliament', 'estate', 'knighthood' - are lodged not in the field of study but in the discipline - like 'period', 'renaissance', 'industrial revolution'. But the concepts of ethnography are, literally, 'far-fetched', as compared to the concepts of the historian.

So there arises the problem: can ethnography be democratized? That is, can it be made accessible to critical verification by a
general public, in our case, by educational practitioners. It is this doubt, which I cannot attempt to resolve here, which in part leads to a current distrust of ethnography and social anthropology in the Third World. The feeling is that ethnography endorsed the power of the outsider - both the western perspective of the ethnographer and the interests of the colonial rulers - and did not substantially increase the power and insight of those who were studied.

It is an important problem for educational ethnography to demonstrate how an ethnography of a school (for instance) can serve those who are studied and their peers rather than simply the ethnographic academy. I have tried to glance at this issue through the problem of verification because I believe that the key to its solution lies in couching educational ethnography in terms which make it much more accessible to verification by reference to the professional second records of educational practitioners. This requires, I believe, a considerable modulation of classic ethnography, the key being in the recognition that the key ethnographic concept of culture is being applied to institutions whose main purpose is the systematic transmission of culture in advanced societies. From different vantage points both ethnographers and educators are centrally concerned with culture. The problem is the design of the bridge which needs to be built between them to establish equity of verification.

Descriptive Studies of Educational Process and Educational Institutions.

Cronbach's plea for more humanistic studies of educational process, educational institutions and educational contexts has been answered and we are now accumulating naturalistic
descriptive studies, both in the United States and in Britain. In Britain these studies range from process studies such as those of Hamilton (1973) and Willis (1979) to institutional studies such as those of Hargreaves (1966) and Lacey (1970). American examples are Smith and Geoffreys (1968), Smith & Keith (1971); Wolcott (1973) and Peshkin (1978). Others could be named.

Behind these published studies is growing a body of dissertations at master's and doctoral levels which handle educational phenomena descriptively, usually studying a single case, occasionally comparing two or a range of cases.

Both as a researcher wishing to use his colleagues' work and as an examiner concerned to appraise the work of candidates I became concerned about the principles of verification of such descriptive studies which are usually based upon tape-recorded interviews, some element of participant or non-participant observation, and usually some documentary sources.

As I have remarked already, verification is not in my view founded upon some direct correspondence between a representation and a reality but rather upon standards which make verification demonstrable, because debatable, in a community - in this case an academic community. The problem is: how shall we ground a discussion of the trustworthiness and coherence of a naturalistic descriptive study?

First, I believe that we must assert that any such study is, as it is written, retrospective. It has no primary source other than notebooks, interview data, documents, and the recollection of the fieldworker, and that recollection needs to be disciplined rather than stimulated by the record (Berk 1977).
That record accordingly should be made available in the public domain in the form in which it was used by the writer of the study: it constitutes the primary source or first record of the naturalistic study. And the existence of microfiche and tape copying techniques make this placing of the record in the public domain technically feasible, as it was not until recently.

The critique of such records rests upon standards which cannot be laid down here and now, but are necessarily created by the community which tends them. However, we can borrow and learn a lot from historians who faced just such a problem when, for example, diplomatic archives were first placed in the public domain in the nineteenth century.

And that itself reminds us that the demand that the records be placed in the public domain will create an archive, will allow not only the verification of naturalistic studies but also their cumulation, will demand therefore that they be indexed. And familiarity with that archive will shape a professional second record alongside the professional second record of direct experience of schools and the commonsense second record.

Thus the research reader can bring a second record at the levels of commonsense, of practical professional experience as an educator and of research experience of the archive to bear on the first record of a particular set of fieldwork notes, interviews and documents, both as a basis for criticizing the naturalistic study and as a basis for generalization across cases. If we do not place primary sources of the kind we are considering in the public domain we do not merely cripple verification and cumulation. We also rob future historians of education of the sources which will allow them to tell the story of education in the intimacy of its process.
REFERENCES


STENHOUSE, Lawrence (1979) The problem of standards in illuminative research. Scottish Educational Review 11,1,5-10
