Evidence and Proof in Documentary Research: I – Some Specific Problems of Documentary Research

Jennifer Platt

Documentary research is not a clearcut and well-recognized category, like survey research or participant observation, in sociological method. It can hardly be regarded as constituting a method, since to say that one will use documents is to say nothing about how one will use them. It is possible, however, that the mere fact of using documents as a data source does pose distinctive problems; the extent to which this is so is one of the themes of this paper.

Discussions of the use of documents in the standard methodological literature are sparse and patchy. In 18 general textbooks on research methods, only 7 devote a significant amount of space to anything to do with the use of documents, and these often either conflate it with other points (e.g. under the general heading of 'unobtrusive measures' or 'available data') or concentrate on only one type of use. Where there are discussions, they tend to be about what types of document exist and what problems they bear on rather than about how to use them; the tacit assumption is made that later chapters, with titles such as 'Analysis of Data' or 'Tests of Hypotheses', deal with that, although their contents usually imply survey-type quantitative data that would only be likely to approximated by content analysis from the main types of documentary research. In the sociological literature there are also very few more specialized monographic discussions of problems of documentary research; indeed I have been unable to identify any since the classic but, I suspect, now largely unread.

work of the late thirties and forties on ‘personal documents’ apart from the literature of content analysis and discussions of official or unofficial statistics and archives. Valuable though such discussions are, they are far from covering all the problems and it seems highly desirable that, at a time of increasing scepticism about the methods with which the textbooks have been traditionally more centrally concerned, some more systematic work should be done in this area. There is some danger that, as part of the reaction against survey methods, any careful concern with method and the justification of conclusions will be abandoned. Insofar as that happens, different research traditions are likely to become closed subcultures with no basis of communication with each other, and even the possible invocation of the concept of paradigm would not make that a happy outcome. The participant-observation / fieldwork tradition has now produced some very valuable systematic accounts of method to set beside those on survey method, and there is no reason why the same should not be done for other areas.

General knowledge suggests that there are some quite different but fairly clear intellectual traditions within sociology associated with reliance on documentary sources: (i) quantitative content analysis, (ii) historical study, (iii) sociology of literature, (iv) linguistic/ethno-methodological, (v) ‘personal documents’. There are also more peripheral uses of documents in research that fit into none of these traditions. This discussion will attempt to concentrate on those aspects of the use of documents which are not specific to any one type of use or tradition; to do otherwise would be in effect either to exclude some major actual uses of documents or to create an ideal type of ‘documentary research’ which included features not logically linked to the use of documents per se. My personal interest in this topic arises from the accidental coming together of several different research topics and experiences, each raising questions intrinsically difficult to answer or not covered by the usual books on method. The main such experience has been work on the history of the antique trade and antique collecting, which entails some conventional historical research, as well as opening up the possibility of some systematic study of the large number of books that have been written about antiques; it also, given the connections of the topic with many aspects of social, economic and artistic history, gives rise to some acute problems in the use of secondary sources. Also relevant have been the problems of analysis in a study based on intensive unstructured interviewing where the interview protocols become documents, and another study using journal articles and textbooks as its basic data. Examples from these research areas will be drawn on to illustrate points, although issues will be discussed in a more general way.

Although the distinction is somewhat arbitrary, for convenience of presentation this first paper will concentrate on those aspects of documentary research which are most directly related to the specific characteristics of documents, while the second will deal with those more obviously shared with other types of research. The precise definition of ‘document’ is not very important, but for the purposes of this discussion it is taken to include at least some of the representational features of pictures without words. (In principle some of the discussion is equally applicable to other kinds of artifacts, and archaeologists have considered some of the issues.) The specific characteristics of documents which raise the greatest problems are that the available stock normally exists already and cannot be created to order, and that contextual information needed to establish their meaning is often missing. For the purposes of this discussion, which concentrates on questions of evidence and proof, the issues arising are defined as being: (i) how to establish the authenticity of a document; (ii) whether the relevant documents are available; (iii) problems of sampling; (iv) how to establish the extent to which a document can be taken to tell the truth about what it describes; (v) how to decide what inferences can be made from a document about matters other than the truth of its factual assertions. Not all of these, of course, necessarily arise or constitute problems in relation to any given research topic or style.

Establishing Authenticity

There are many instances where a document is not an authentic member of the class to which it purports or appears to belong: wills, other legal documents and letters are forged or falsified, literary works are attributed to authors who did not write them, jokes and satires may inadvertently be taken at face value. An inauthentic document may still be of much interest, but it cannot be fully and correctly understood unless one knows that it is not authentic. Procedures are, therefore, needed for establishing authenticity. Criteria are also, in principle, needed for deciding when it is worth applying such procedures; in practice the normal criterion is only to look closely when there is felt to be some reason for suspicion in a particular case, or when in-authentic examples of the class are known to be common. Guidelines would still be desirable to indicate when there should be felt to be reason for suspicion: the simplest case is that where the sociologist does not need to decide because authenticity is already disputed among those studied, as in many legal cases. I have not come across any systematic list of other suspicious circumstances, but among those referred to in passing are the following, almost all of which are equally relevant when the question is not whether the document as a whole is authentic but whether the version one has of it is correct or has errors arising from copying, reproduction or other modes of transmission: (i) the document as it stands does not appear to make sense, or has obvious errors in it; (ii) different versions of the same original document are current; (iii) the document contains internal inconsistencies of literary style, content, typeface or handwriting, etc; (iv) the document is known to have been transmitted via many copiers; (v) the document is known to have been transmitted via someone with an intellectual or material interest in the version given passing as
the correct one; (vi) the version available is derived from a secondary source suspected on other grounds of being unreliable; (vii) the style or content are in some way inconsistent with that of other instances of the same class – e.g., it contains anachronisms; (viii) it fits too neatly into a standard formula or literary form. It will be noted that most of these criteria can only be applied if one already has some other relevant knowledge of the class of document and/or its social context, and that this ‘knowledge’ in turn may be grounded on documentary sources which are in principle equally open to question.

Once the question of authenticity has been raised, an attempt must be made to answer it. This may be done either by an analysis of the document itself or by exploring other circumstances associated with it. Historians (and lawyers) have developed elaborate methods of detection involving such matters as the types of paper and styles of handwriting typical of periods or individuals, the characteristic kinds of error made in careless copying of different kinds, and so on; art historians lay tremendous emphasis on known provenance of works of art as a guide to their authenticity – although provenances have in their turn been falsified. Sociologists seldom in practice seem to be faced with the problems to which these techniques can provide solutions, since they do not often work with primary sources which are handwritten or drawn and/or regarded as individually important by people in a position to tamper with them without immediate detection.

Availability of Documents

The next issue is that of the availability of suitable documents. Unless the documents used are generated by the sociologist, which is inappropriate or impossible in many fields, he must use what he can get. Some documents one might like to have will never have existed, others have been lost or destroyed, and others still exist but one cannot get access to them. The equivalent of the last case could arise with any type of data, so will not be considered here. The loss and destruction of documents gives rise to two kinds of difficulty: an inadequate quantity of data, and, where there is reason to believe that there are differential loss rates, a qualitatively unsatisfactory distribution of data. Where there are simply not enough data, one is tempted to over-interpret what is available, and to treat it as representative of the larger class that originally existed without any knowledge that it really is so. Without knowledge of other members of the class one cannot know what the conventions of the genre are, or indeed whether there is a defined genre and if so where its boundaries lie. A single reference to a phenomenon may indicate the start of a trend, or the existence of a pattern, but may be just historically idiosyncratic. Where there have been differential loss rates this makes what survives less representative, although the knowledge that it is so gives a firmer basis for interpreting what does survive. Often, however, one does not know whether the loss rate was differential or not; one may have reason to believe that it was, but this ‘reason’ may be to some extent speculative rather than grounded in specific data. For instance, it seems plausible that those few cases where the accounts of early antique dealers have survived to the present day should be representative of the whole body of dealers existing at the time, partly simply because so few have survived, and partly because one reason for survival of accounts is the survival of the business where others have not survived. Unless, however, one has other sets of accounts, or other sources of data on the typicality of the business, one cannot be sure. Even if one approaches the matter conservatively, and treats the accounts as those of an unusually solid and successful business, this does not tell one how a successful business at that time differed from an unsuccessful one – although it may refuse some hypotheses about the necessary conditions for success. There are some cases where the surviving documents may be regarded, not as typical of the original population, but as more significant, because their survival indicates that they were more popular or more highly valued. Unsuccessful books printed in small editions are less likely to be found now than very popular ones. This is a great consolation to the frequenter of secondhand bookshops, but one of limited value, since it does not help if one is interested in what was originally produced rather than what became successful, or in ephemera on cheap paper as much as calfskin editions with prestigious illustrations.

The case of the documents that never did exist raises problems analogous to those of perspective in fieldwork. Very often only one set of parties to a relationship produces documents that bear on it, and such documents are obviously likely to be misleading about the total of what went on. One is tempted to make inferences on the basis of intuition, general knowledge, or theory about what the content of the ‘missing’ documents would have been, but this is clearly dangerous. It may also happen that those who write are, rather than representative of one set of parties to a relationship where the other side is unrepresented, simply an unrepresentative sub-set of the larger group. One might argue that sufficient senior politicians write their memoirs for these memoirs to be regarded as representative of senior politicians more generally, especially in view of the tendency for opponents each to try to justify their actions. But there seems every reason to believe that those antique dealers who write their memoirs are in that respect atypical of antique dealers more generally. What general meaning, then, can one attach to what they say? One can certainly go beyond just treating them as idiosyncratic. If most of the authors describe their background as middle class and explain how they were drawn to the trade by their interest in collecting beautiful things, it would be rash indeed to assume that this was the typical pattern, in particular because it seems likely that such characteristics would make dealers more prone to write books. But given that one knows that, say, size of business, character of town and clientele, period of specialization and type of goods dealt in are significant distinctions, one can check the distribution of the businesses reported
on with respect to these variables; if one finds a wide variety, and/or if the same points are reported irrespective of such variations, one has greater confidence that what is described can be taken as representative.

However, there are certain types of anecdote that recur again and again; one cannot infer from this that such situations occurred in life more frequently than others which do not make such good anecdotes, although the fact that anecdotes of that type appear in many books may suggest that they describe common experiences of dealers. Another possible line of argument is that if most of the books express views that run counter to the norms of the wider community, or offer matters sufficiently technical for the wider community to have no norms on them, then they must be regarded as at least to some extent typical of antique dealers as opposed to the general population — though how typical is unclear. Again, if the books — as they do — take opposed stands on a controversial and professionally important matter such as auction rings, this evidently demonstrates that they are not reporting only one side of an argument, though it tells one nothing about the prevalence of the different viewpoints or the possible existence of yet other ones. If, however, among those who approve of rings a rather uniform set of justifications for them appears, this could demonstrate not so much a shared subculture as a shared understanding about what is likely to seem persuasive to the readers of such books; perhaps it makes the imputation of a subculture more plausible if the norms stated in overt argument are also demonstrated in action in anecdotes whose central point is a different one.

Convincing as these suggestions of criteria may seem, it must again be noted that they rest on commonsense assumptions on such matters as who writes books and how, and at best allow for greater confidence rather than secure inference.

Sampling Problems

These considerations are very much related to questions of sampling; indeed, they may be regarded as indicating where the problem is one of defining the sample one has got rather than of choosing what sample to select. But sometimes in documentary research there are large numbers of potentially relevant documents, and the possibility of sampling in the usual sense arises. An initial difficulty, when one cannot create more data (by further observation, conducting interviews, etc.) is that there will be no fresh data on which one’s hypotheses can be tested; this suggests that, if one wishes to employ a hypothesis-testing model of the research process, one must be careful not to use all the data up in exploratory stages, and this might suggest some sampling procedure, but useful samples can only be taken if one knows something about the population and has a sampling frame. But unless other researchers have been there first one may well not know enough about the population, or have anything that could serve as a sampling frame.

G. R. Elton suggests that, where no lists exist, the historian should deal with the problem of abundance by controlled selection, that is to say by choosing on principles which have nothing to do with the real question asked or the ultimate product . . . the first selection . . . should arise from a total survey of the material and be systematic with reference to it, not to the historian’s purpose. This suggestion is intended to prevent the bias that might arise from selecting only those documents that bore on, or even supported, a particular preconceived hypothesis; obviously it implies a preference for research where the questions asked arise from the data rather than from prior theoretical interests, and assumes that the historian has already decided enough about what he is substantively to study to select a particular archive on which to start. I take it that the arguments about the impossibility or undesirability of the approach are too familiar to need rehearsing here. In this case, as in others, a very strong emphasis on faithfulness to the specificity of the data entails a theoretically arbitrary delimitation of which data shall actually be collected, at least initially. Contemporary social groups and organizations often have convenient physical boundaries, and it is noticable that fieldworkers normally choose such settings; the archive probably represents the nearest documentary equivalent. Since data are not always conveniently arranged into archives, even to accept Elton’s principle would often leave one with little guidance.

Whether or not one follows a hypothesis-testing model, the problem of creating a sampling frame may still arise. For some categories of documents, comprehensive listings do exist, whether, like catalogues of Roman inscriptions or some bibliographies, they have been devised for research purposes or, like catalogues of books in print or lists of wills, they have not. Whatever the purpose for which they have been devised, it is highly probable that the categories of documents have been defined more narrowly or broadly than is suitable for a new researcher’s purposes, or that the principles of classification cut across those which he would wish to use. For instance, in relation to my research on antiques I was delighted to find a comprehensive bibliography on glass, but became somewhat less delighted when it turned out to include every aspect of the technology of glass manufacture, and works on industrial and commercial uses of glass, as well as those aspects which interested me.

One may thus be driven, as I have been, to attempt to compile one’s own comprehensive listing of documents in the relevant field: this is an enormously laborious task. Titles or descriptions can be misleading, so only if one inspects each document can one be quite sure that it is of the type one is interested in; The China Hunters’ Club, by ‘The Youngest Member,’ turns out to be an amusing fiction rather than the true history of an early collectors’ club. Many sources of information are careless, or follow unfamiliar conventions, which can lead to misunderstandings; I have often found that what appeared to be
two separate books on antiques was actually the same book listed in different ways in different sources. The subject catalogue of the National Art Library at the Victoria and Albert Museum is splendidly comprehensive, indeed excessively so for some purposes; every entry in some parts of it has to be checked against the author catalogue to find whether it refers to a book or only to a brief article. Workers on other topics will certainly be able to add to these examples. A related problem is how to know when to stop; how many defunct journals’ review sections, secondhand book catalogues or publishers’ advertisements in the backs of old books must one read before being entitled to conclude that a further search is not likely to turn up anything new? After spending some weeks on this sort of work, one may wonder whether one’s research time could not be better spent in other ways.

The question of when to stop is not a purely rhetorical one. The answer to it is easier if the research is mainly concerned with documents that have been influential, for these, by definition, will have left traces behind them; the search does not need to be extended beyond a reasonable cross-section of the documents in principle likely to have been influenced. (Note, however, the assumption here that the relevant influence is one that is primarily shown in documents; even when one is studying the intellectual history of an academic discipline, there may be reason to believe that there are components of the oral subculture not adequately represented in what is published.) If, however, the research really is concerned with the whole range of documents which originally existed (highly desirable if Whig versions of history are to be avoided) the answer to the question of when to stop must be something more like ‘when all potential types of source have been thoroughly sampled and several further instances of each type do not bring anything new to light’.

(This assumes good knowledge of available types of source, which is easier to come by in fields already worked over by others.) Pragmatic and/or principled judgments may also be made, when time or other resources are limited, about how far the likely return justifies further input on this rather than other aspects of the research. But can such tasks of enumeration be avoided? In practice one may suspect that they frequently are; in principle it is hard to see any alternative if one wishes in the end to make well-founded statements about, say, the number and character of books on antiques published at various times, or changes in the types of distinction among research methods made by sociologists, let alone to draw samples for more detailed study which can claim to be representative.

A quite different problem which arises in the sampling of documents is that of whether they are independent of each other. I was pleased to find two books by the same authors on much the same topic 35 years apart, thinking that this would give a good indication of changes over the period; on reading the second, however, I found that a considerable number of the chapters were identical.

A commoner problem is that of plagiarism of the works of others. Antiques have come to be studied in an increasingly scholarly way, and there are many complex details involved in questions of attribution, identification and dating; there is also much popular interest of an unscholarly kind. In consequence there are many popularizing books which draw on the scholarly work, more or less competently. The fine between plagiarism and popularization is hard to draw, depending to a considerable extent on the degree of acknowledgment given to the work of other authors; certainly the scholarly writers have sometimes seen the popular work as both plagiaristic and incompetent. Where there is plagiarism, or for that matter acknowledged copying, the work may still be treated as an independent case for some purposes, but not for others. Whether or not there is any element of plagiarism, similar issues may arise where there are clearly-defined genres; the writing of a 14-line poem with a certain kind of rhyme-scheme cannot be treated as a personal fact about the poet in the same way as the writing of a poem of non-standardized form. (For further discussion of the question of genre, see the second part of this paper.)

Does It Tell the Truth?

The documentary researcher is often using the document as a source of information on to which it refers – in effect as a surrogate researcher; it is vital, then, to know how far its account can be relied upon. This is a matter of how the document is used rather than of what it is in itself; a legal document may be studied as an example of legal forms, or as an account of a particular contract; a picture may be looked at as a work of art in its own right, or studied for the light it throws on costume and furniture when it was painted; a work by another social scientist may be treated as a datum relevant to conventions on how to present social-scientific data, or as a source of data on the matters about which it writes. When a document is being used in the second way, the user must consider how far it can be taken to tell the truth about what it describes. What criteria can we apply to help us make such judgments?

Probably one of the first writers to deal with this subject systematically was John Craig. He starts from some simple assumptions: that ‘All men have an equal right to be believed, unless the contrary has been established from elsewhere’, so ‘great probability is composed of many testimonies of primary witnesses’; the probability of a report is decreased in proportion to the distance in time or place of the event and the number of people through whom it has been transmitted. By adding a few numerical assumptions he is enabled to conclude, by Newtonian method, that the present probability of the story of Christ is as great as that which a man (in the times of Christ himself) would have had who received the history by only word of mouth from 28 disciples of Christ. Q.E.D. This is an attractive approach, though unfortunately it has some serious difficulties both in the particular empirical assumptions that it makes and in its failure to consider the possibility that one may not know through how many people a story has been transmitted.
Craig's emphasis on the importance of direct witnesses, and the attempt to evaluate the degree of proximity of a report to them, have also been taken up by other, later writers. Naroll suggests a ranking of types of secondary source by degree of proximity, in this order: datum report, participant report, observer's report, derivative report, scholar's report, (citing primary sources), reader's report (not citing specific passages from primary sources). He also suggests six factors relevant for data quality control where the data are reports on practices in alien societies: (i) collection of specific case reports, (ii) use of direct observation and participation, (iii) length of stay in the society, (iv) reporter's familiarity with the language, (v) the reporter's role in the society, (vi) the explicitness and generality of the report on the matter of interest (p. 14). These factors are all concerned with either how full and detailed the information was to which the reporter had access, or with how much care and detail he reported on it. The list was compiled in the attempt to answer the problem of which author to believe when there were conflicting reports on the same society, but it could to some extent be used to judge how far to accept a single report. He suggests yet another list of possible criteria in relation to historical data: (i) the time lapse between the event and the report, (ii) the extent to which the author had a professional stake in the report, (iii) the extent of agreement between authors with opposing views, (iv) the proximity of the author to the event recorded, (v) the intensity of involvement of the author with the event, (vi) the degree of relevant technical experience of the author and (vii) the degree of explicitness of the report. It will be noted that this list for the first time raises the possibility of reporting distorted by personal interests rather than just by lack of full information about what happened, and so introduces a whole range of other questions.

These questions are addressed directly by Langlois & Seignobos, who list a variety of possible sources of distortion which might affect one's general or particular confidence in an author. Among those that add something to those quoted above are these: (i) that the author was seeking some practical advantage to himself; (ii) that he was in a situation where he was compelled to violate the truth, as when, for example, drawing up a formal document about a situation where practice was inconsistent with rule or custom; (iii) his vanity led to exaltation of himself or his group; (iv) he wanted to please his public, or to avoid shocking it; (v) he tried to please by literary artifice, introducing misleading elements of the picturesque, noble or dramatic; (vi) although in a position to observe the events reported, he was led by illusion or prejudice to observe incorrectly; (vii) he was badly situated for correct observation; (viii) he could observe, but did not record his observations immediately in a full and unambiguous system of notation; (ix) he did not actually observe the events, but gives an officially-correct or customary version of them; (x) the facts reported are such that they could not have been learnt by observation alone, because they were normally secret or private, or are generalizations about large numbers of people over long periods of time. As the authors point out, these apply where the statements are, or purported to be, based on firsthand observation; where they are secondhand, they should in principle be applied to the first observer as well.

It is evident that all these possible sources of distortion may indeed have led to biased reports. As one modern sociological textbook remarks, 'With answers to these questions, the investigator will usually be able to eliminate records that are in doubt and accept the rest as credible evidence.' None of the textbooks, however, provide much explicit guidance on how one is to get answers to such questions, which is the practical problem with which the documentary researcher is faced. Nor do they raise any questions about the existence of a simple, factual truth in relation to which accounts can be judged as biased, or tell the investigator what to do if all the available documents appear to be biased in one way or another. (Except that the implication of the quotation from Nachmias & Nachmias is that what he should do is give up; this may be true, but is not helpful.)

Some of the questions seem rather easier to answer than others, at least when the circumstances are favorable. If the author has made his own position and methods of data collection clear, one may judge their implications: if he has not, another contemporary source may provide the same sort of information about the author in person, or about general practice at the time. However, it is not always clear what one should do with such information: someone who was not in a very good position to see may nonetheless happen to have seen the key events correctly, and even a royalist author of literary bent may be reporting nothing but the truth, as even a cynical republican would recognize it, when he describes the chivalrous behaviour of the king. Besides, there is an obvious danger of circularity, inferring royalism from the favourable picture given of the king and then using the inference to devalue the account. Where more than one account of the same events is available, the matter becomes different: one may legitimately infer that the differences are attributable to difference of viewpoint, physical or intellectual, and attempt to construct from the sources a superior composite account. But such a composite account is not necessarily superior; one source may simply be more accurate and the other not. If, other things being equal, one source had better access than the other, a superior recording technique, or less personal involvement, one presumably accepts, or at any rate prefers, its account and rejects the other. But the circumstances are not always favourable; authors frequently do not describe their situation and methods, and other accounts of the same events are frequently not available.

If there is only one source, one may attempt to make adjustments for its biases, assuming that the king was probably not quite so consistently chivalrous, the speech not exactly in the words of the supposedly verbatim account, the role played by the author's group not quite so essential or conspicuous; this is inevitably a speculative and uncertain activity. It also takes for granted some
commonsense theories about the processes involved. In principle these theories might themselves be tested, so that the typical errors of reporting could be identified as precisely as the typical errors made in copying from different types of handwriting have been. I do not know of work specifically directed to this aim, except perhaps some by psychologists on perception and memory; Bartlett’s classic work is highly suggestive. But it seems very likely that the typical specific errors made by different groups at various times and places are not the same, so modern research which provides the only way in which a privileged set of observations could be constructed with which to compare reports, might not be much help in relation to historical documents.

An alternative or supplementary approach to the single source is to consider the likelihood of distortion in relation to the subject matter. Langlois & Seignobos suggest that there are some sorts of fact about which it is difficult to lie or to be mistaken, because they are such that they are very easy to observe (e.g. the existence of a city) or were widely known at the time among the audience to which the writing was directed. They also point out that one’s confidence can reasonably be greater when the author was not particularly interested in a specific fact and mentions it only incidentally or when it runs counter to his expectations and habits of mind or the general impression he wished to produce. Once again the points are convincing, but they both assume the absence of subtle forms of bad faith which have anticipated the application of such criteria, and involve commonsense theories and the danger of circularity.

I do not wish to discuss here the senses in which there may be taken to be a simple, factual truth of which various reports give more or less satisfactory accounts. The absence of epistemological discussion from the writings of historians on these matters is indeed refreshing, if at times somewhat startling; how important such absence is depends on the nature of the accounts and of the information that is wanted from them. To treat all accounts as merely ‘accounts’ solves some problems but, like operationalism in relation to problems of validity, does so at the cost of ignoring the reason for being interested in the matter in the first place.

But if, on whatever grounds, all the available documents appear to some extent biased, what is one to do about it? It is not feasible, except possibly in very unusual circumstances, simply to subtract the erroneous parts and to proceed securely on the basis of the remainder, both because of the likelihood that there would be no remainder, and because even if probable error can be identified it is not always thereby evident what would be correct. It has been suggested in the literature that there are some occasions when independent but separately dubious sources can give mutual support either by conveying the same general impression or by providing different factual items that fit together into a coherent picture of one event, and this is a valuable suggestion, despite the elements of judgment implied by the concepts of sameness and coherence. Beyond that, one is thrown back by some writers on such criteria as ‘historical imagination’ or that ‘answers must be probable: they must agree with what is known to be possible in human experience’ this means, in effect, the application of commonsense theories.

Todd has a much more satisfactory and more explicit approach, which he works out by examining instances of the way historians reason in practice. He quotes an example from Bullock’s work, where he decides to believe a document from one participant in an episode rather than another: ‘... there were basically two Papens that fit the evidence, and... they are such that we can understand either one... (but) if we suppose that Schröder’s was lying and Papen was telling the truth we cannot see how to put ourselves in Schröder’s shoes at all, at least without inventing circumstances for which there is no evidence. But on the hypothesis that Schröder was telling the truth and Papen lying we can imaginatively re-enact both their roles. Hence the latter hypothesis is to be preferred.’ This approach can be highly convincing in a context of rich data, where there is other evidence besides that of conflicting documents to help one interpret them and where there are many details which an interpretative hypothesis must fit and which therefore might refute it; where data were thin, it could be dangerously speculative. Even where there are not two documents with conflicting accounts of the same event, but one document and some contextual information, the same strategy may be used (and probably often is) to make judgments about the extent to which it can be trusted. As Todd points out, an alternative approach is to apply supposed psychological laws, e.g. ‘Witnesses never distort the truth in a way unfavourable to themselves’. Even if true in general, such laws ought only to be applied if the hypothesis that they exist has been directly investigated, but it would scarcely be possible for the researcher on a particular topic to do such research himself as it turned out to be relevant. However, it might be more convincing, as well as more feasible, if it could be shown that there were customs and intellectual styles of this sort specific to the milieu producing the documents of interest.

What Inferences Can Be Made?

These considerations also bear on the next aspect of our topic: the making of inferences from documents about matters which they do not directly describe. The most minimal inference occurs when the existence of conventions about the form which documents of a certain kind should take is inferred from the presence of certain kinds of similarity among them. It is of the utmost importance for interpretation that such inferences be made, though in the absence of supporting information (e.g. textbooks which give instruction on the writing of correct sonnets, journal articles or legal contracts) it is not obvious how to distinguish between the effects of a convention and of the same causes independently producing the same effects. (A study was done in
which respondents were asked to write convincing suicide notes, and the results were then compared with genuine suicide notes and found to have significant differences from them.35 Such data suggest, on the assumption that the sample of respondents had equal access to the conventions on how to write suicide notes, that what the authentic notes had in common was not conventional. Data of this type are desirable, though seldom likely to be available.) Becker has argued, to me convincingly, that the conventions about how to write journal articles ensure that they are unrepresentative of all the sociology that is done36 and other writers have made similar points.37 It is curious, in the light of this, that in the large number of citation studies it is almost invariably assumed that frequency of citations can be taken as an indication of the influence or prestige of the work cited, and the possible influence of convention is not considered.

This inference is characteristic of the inferences that may be made from the content of documents to the beliefs, motives etc. of those producing them; the connecting link is a theory about the reasons why such beliefs etc. should be expected to produce documents with such content. The inference can be justified to the extent that the theory is known to be correct. My comment above about citation studies has two implications: (i) that independent evidence is needed to support such a theory, and to get such evidence one would need to investigate the theory itself, and (ii) the inference depending on such a theory is weakened further to the extent that there is an alternative theory which could also account for the form the document takes. It follows from this latter, however, that if no such alternative theory can be found the inference has some plausibility even in the absence of independent support for the theory.38 Such inferences have been made both to the characteristics of groups and to those of authors as individuals. Whether such inferences seem plausible will depend not only on the support available for connecting theories, but also on sampling considerations: if only a few of Shakespeare’s plays survived, one might draw conclusions from their characteristic imagery about his personality; given that many survive, and show different patterns of imagery, one is more likely to draw conclusions about his art.39 Letters, diaries and autobiographies may give a graphic picture of the characters and attitudes of some individuals; whether they can be taken as representative of their class or time is another matter. (It could well be argued that when one is most tempted to make such an assumption, because of the absence of other data, it is most likely that those who have produced such documents are thereby unrepresentative. This too is, of course, a connecting theory, whose merits can be investigated.)

On other occasions the desired inference is not to the beliefs etc. of the author but to the real state of affairs lying behind the production of the document. There are numbers of examples of analyses of propaganda36 where, for instance, the real state of German/Italian relations or of successes in battle on a particular front, is inferred from statements not overtly about these facts or even, perhaps, suggesting the opposite. Here what seems to be happening is that it is the motives of the authors that are taken as given, and the inference made is about the documentary content that such motives would lead to in varying objective situations. The inference might be grounded on previous data about what was said in various situations on which independent data are available, or might be speculative/common sensical: whichever it is, the connecting theory is again necessary.

Another kind of inference commonly attempted is that from the document to the characteristics of its audience. What the document directly shows, more or less implicitly, is its author’s assumptions about his audience – a point which one needs to bear in mind when interpreting it as evidence of his personal concerns; these must have some weight, since he is likely to have been more familiar with them than the researcher is. However, to address oneself to an audience is not the same as to be identified with it, although one may feel that in a field such as that of antique collecting, where the boundary between amateur and professional is vague and often crossed, or writing sociology, where writers and audience are to a large extent the same group, writers can more easily be assumed to be representative of their audience. Different kinds of work show varying degrees of consciousness of their audience – some seek popularity, others aim to manipulate, or are pure self-expression, and so on. Once again evidence is needed to identify what is going on. Relevant types of evidence are either independent documentation (sales figures, reviews, comments by contemporaries on reception, audience surveys) or confirmations of the connecting theories needed to support inferences from content to audience.

Conclusion

A brief review of the main topics covered in this paper will consider how far each reveals problems distinctive of documentary research. Authenticity, in the sense where the alternative is forgery or mistaken authorship, can only apply to documents; as soon, however, as the sense is slightly broadened to cover other forms of incorrect identification or deliberate misrepresentation or use of false fronts the essential problems are clearly the same as when the data are observations or reports of other kinds.40 The questions of existence and availability of desired documents have much in common with practical issues of access, perspective and social visibility in observational research. The status of different types of account and their recurring patterns, and how to evaluate it, is a quite general problem; the documentary researcher merely does not have the opportunity of the fieldworker to generate fresh data of the same or other kinds to get behind the original accounts.41 unless those who produced the documents are still alive. The problems of lack of choice in sampling, and the need to construct one’s own sampling frame, arise in many
other types of research (even if they are not always confronted); documents have less propensity to move house or fail to co-operate than people, so may make the task easier. Methods of search for individuals in an unknown population of documents, and of deciding when to stop searching, seem closely analogous to issues raised by 'snowball' samples for interviewing. The making of inferences from documents is in essence much like the making of inferences from other forms of behaviour, with the practical limitation again that there is likely to be less possibility of laying hands on supplementary information desirable to assist inference.

We conclude, therefore, that there are important senses in which documentary research has problems which are not significantly different from those of research using other data sources. A more general moral might be suggested, which is that the distinctions commonly made among 'methods' of research in terms of their data sources may be analytically unhelpful. (It is hoped to pursue this point further elsewhere.) Nonetheless, each data source gives rise to its own special technical problems. If adequate 'connecting theories' are to be developed, the form which they will take must be one of specifying, firstly, general propositions about how people behave (other things being equal, people act in ways they construe as being to the advantage of their families); given such propositions, their application to the particular type of data to be used must be developed ('in societies where suicide is seen as shameful, the suicides of those with surviving close relatives will be under-reported'); ideally, they would then be elaborated into precise technical recipes for dealing with the immediate data ('in society x, the actual suicide rate should be estimated as y% higher than that given in the official statistics'). Such propositions are empirical ones exactly like those in sociology generally: method and methodology cannot be separated from substantive concerns — and it follows that substantive research is needed for them to progress.

Notes

The support of a Leverhulme Foundation Research Fellowship during the period when much of the work for this paper was done is gratefully acknowledged.


2. Peter Mann's Methods of Sociological Inquiry (Blackwell, Oxford, 1968) is an honourable exception to this generalization, although it does not touch on analysis.


11. A classic example of the problems that can arise for social science from poor reproduction is Freud's analysis of Leonardo da Vinci's sexual attitudes on the basis of an incorrect reproduction of a detail in one of his drawings.


T
c

The Immigrant Letter between Positivism and Populism: The Uses of Immigrant Personal Correspondence in Twentieth-century American Scholarship

David A. Gerber

To justify leaving home, to establish migration chains and teach those about to join them about the experience of emigration, to ask for or to provide economic resources, to attain intimacy, and to maintain emotional bonds, immigrants before the era of instant electronic communications were compelled to write letters to family and friends in their homelands, even though their literacy skills were often quite rudimentary. The result is a vast, unique archive of the writings of obscure and ordinary people of dozens of different ethnic-cultural backgrounds that is especially rich for the nineteenth and early twentieth centuries, during which the routinization of transoceanic shipping facilitated the exchange of mail between Europe and North America. To be sure, not all immigrants participated in this international exchange of personal letters: illiterates; those who had severed family ties; those with completed families; children; and women (who even when literate were often spoken for in correspondence by husbands, fathers, or brothers), all are underrepresented among letter-writers. We cannot find all immigrants even in this immense archive. That is certainly not the end of the fundamental difficulties the letter poses. In many instances it is difficult to learn anything beyond basic biographical facts about either letter-writers or those to whom they wrote. Only occasionally, moreover, do we have access to the letters that were sent to our