

# A MATTER OF RECORD



J O H N S C O T T

# Contents

## List of Figures

## Preface

## 1 Social Research and Documentary Sources

Evidence and data in social research

What are documents?

## 2 Assessing Documentary Sources

Authenticity: soundness and authorship

Credibility: sincerity and accuracy

Representativeness: survival and availability

Meaning: literal and interpretative understanding

## 3 The Use of Documents in Social Research

The search for King Arthur

Who wrote the Zinoviev letter?

The social meanings of suicide

Social research and the relativity of accounts

## 4 The Official Realm: Public and Private

The State, surveillance and secrecy

Official documents in the State

Official records in the private sphere

## 5 Administrative Routines and Situated Decisions

Conceptual instruments and administrative routines

Situated decisions

## 6 Explorations in Official Documents

Occupation, class and inequality

Class schemas and the problem of meaning

Measuring the class structure

The records of health, welfare and education

Business and industrial records

## 7 The Public Sphere and Mass Communication

Public opinion, the media and the audience

Images, content and meaning

Directories, almanacs and yearbooks

The annals of the rich and powerful

## 8 Personal Documents

Diaries, letters and autobiographies

Photographs and visual sources

Notes

Index

# A MATTER OF RECORD

Documentary Sources in Social Research

*John Scott*

Polity Press

Copyright © 1990 John Scott

First published 1990 by Polity Press in association with Basil Blackwell

Editorial office:

Polity Press

65 Bridge Street,  
Cambridge CB2 1UR, UK

Marketing and production:

Basil Blackwell Ltd

108 Cowley Road, Oxford OX4 1JF, UK

Basil Blackwell Inc.

3 Cambridge Center

Cambridge, MA 02142, USA

All rights reserved. Except for the quotation of short passages for the purposes of criticism and review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

Except in the United States of America, this book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, resold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

ISBN 0 7456 0030 1

ISBN 0 7456 0070 0 (pbk)

*British Library Cataloguing in Publication Data*

A CIP catalogue record for this book is available from the British Library.

*Library of Congress Cataloging in Publication Data*

A CIP catalogue record for this book is available from the Library of Congress.



# Figures

[1.1](#) Sources of evidence: proximate access

[1.2](#) A classification of documents

[5.1](#) The Registrar-General's classifications

[6.1](#) The Hall-Jones classification

[6.2](#) The Goldthorpe classification: version 1

[6.3](#) The Goldthorpe classification: version 2

[7.1](#) Women's magazine covers, 1949-74

[7.2](#) A typology of directories 171

# Preface

Documentary sources of information, of all kinds, figure centrally in the research of sociologists. Official statistics on crime, income distribution, health and illness, censuses of population, newspaper reports, diaries, reference books, government publications, and similar sources are the basis of much social research by academics and their students. Yet these materials have rarely been given the attention that they deserve in accounts of sociological research methods. Questionnaires and participant observation figure centrally in texts and courses on research methods, but documentary sources are considered in only a fragmentary way.

The aim of this book is to attempt to remedy this situation by illustrating the diversity of documentary sources available for social research, and to discuss some of the ways in which they can be used. In doing this, I emphasise the similarity of the methodological problems faced by sociologists, historians, and other social scientists who use these sources in their work. This argument implies certain claims about the drawing of disciplinary boundaries – I am an unashamed ‘sociological imperialist’. But I have deferred these general considerations to another occasion. My focus is on the handling of documents in relation to specific problems in social and historical research.

I also make the claim that the methodological issues involved in handling documentary sources are similar to those that arise in handling any sources of evidence in social research. I introduce four criteria for assessing the quality of social research evidence – authenticity, credibility, representativeness, and meaning – and I outline their application to the whole range of social research before



proceeding to a detailed consideration of documentary sources. These criteria are used throughout the book as organising principles for my discussion. Although it is unnecessary that a researcher use them in the same systematic way, my argument is that an appreciation of their importance can only be learned through considering them systematically.

The examples of research that I have used in the book are drawn from a number of different substantive areas, and I have used both American and British source material and studies. My belief is that the uses and limitations of documentary sources can only be fully appreciated when they are understood in their social context as historical products. To this end, I have concentrated my attention on British source materials and have shown their emergence in the context of the British state and economy. Wherever appropriate, however, I draw parallels with the wider European and American experience and the types of source material that emerged in other countries.

My initial concern to write this book was stimulated by an important article by Jennifer Platt, which is referred to at numerous places in my text. Her path-breaking article not only suggested the need for such a book, but it also advanced some of the ideas that underlie my own development of the four appraisal criteria for social research materials. I am greatly indebted to her for pointing the way. The argument of the book has also been shaped by my involvement in the teaching of research sources and methods over a number of years, and by my own involvement in documentary research. I am grateful to Tony Giddens and Polity Press for the patience that they have shown in waiting for the final manuscript, my production of it being delayed by the administrative pressures of modern academic life and an overloaded mainframe computer.

John Scott

# 1

## **Social Research and Documentary Sources**

The handling of documentary sources – government papers, diaries, newspapers and so on – is widely seen as the hallmark of the professional historian, whereas the sociologist has generally been identified with the use of questionnaires and interview techniques. In fact, documentary investigation was the main research tool of the classical sociologists: Marx made extensive use of the reports of the factory inspectors, Weber utilised religious tracts and pamphlets, and Durkheim employed official statistics on suicide. The bulk of the historical and comparative work that is undertaken in contemporary sociology involves the use of documentary materials, as does much work on contemporary societies. But textbooks on research methods have generally failed to recognise this and have given most of their space to discussions of questionnaires, interviews, and participant observation. The aim of this book is to rectify this imbalance by considering some of the issues involved in dealing with documentary evidence.

My argument is that the general principles involved in handling documents are no different from those involved in any other area of social research, but that the specific features of documentary sources do require the consideration of their distinguishing features and the particular techniques needed to handle them.<sup>1</sup> This argument rests upon the stronger claim that the logic of

social investigation is no different from that employed in those areas of research concerned with physical objects, chemical substances and living bodies. All research involves the systematic and disciplined search for knowledge of the world that exists outside of the researcher's laboratory, institute or department. This is not to say that no methodological peculiarities arise in the study of social phenomena. There are, indeed, features of concept formation and theorisation in the social sciences which derive from the inherently 'meaningful' and 'value laden' character of social reality. These features do not, however, require any departure from a 'scientific' frame of reference. The meaningful character of social reality, moreover, does not constitute a 'problem' for social science methodology. It does mean that it is difficult to achieve the kind of unambiguous descriptions produced in the natural sciences, but the social scientist has certain compensating advantages. The inert phenomena of physical and chemical science are unable to communicate meaning to the scientist in the same way that human beings are able to, through written documents and through conversation. The argument of this book, therefore, is that sociology is a science and that documents should be handled scientifically.

The social researcher has available to him or her an enviable array of source materials from which scientific data can be constructed. The researcher can interview people, observe their behaviour, interact with them, study the administrative records of their dealings with official agencies, read their letters and diaries, examine the newspapers they read, and so on. Each particular source requires slightly different handling: the mechanics of participant observation, for example, are radically different from those involved in the use of official statistics. Written documents of all types, however, share certain features which distinguish them from other kinds of source material.

For this reason, it is necessary to understand written documents in the wide context of the whole gamut of sources which are utilised in social research.

## **Evidence and data in social research**

The aim of social research is to describe and explain the actions of agents and the structures that they produce and reproduce in the course of their lives. But neither 'actions' nor 'structures' are actually observable: they are inferred from the behavioural and other observational evidence through which they are manifested. It is important therefore to examine the types of evidence available to the social scientist as source material for data construction. This can best be approached by considering two contrasting relationships between the observer and the observed.

*Proximate* or 'direct' access by the observer exists where the observer and his or her source material are contemporaneous and co-present. In such a situation the observer is a direct witness of the audible, visible, and tactile signs of human action: the observer and the observed are 'coincident'; they have the same spatio-temporal location, and the observer may therefore use eyes, ears and other senses to observe and question the current behaviour of those being observed. *Mediate* or 'indirect' access by an observer exists where past behaviour must be inferred from its material traces, the visible signs of what happened or existed at some previous time. In this case the observer and the observed are not co-present and the observer may only obtain evidence indirectly from the buildings, books and so forth produced by people in the past. The distinction between proximate and mediate access should not, perhaps, be drawn too sharply, as much

everyday 'observation', for example, involves both processes. We see people move and hear them speak, and we are also able to see the houses and cars they have bought and the clothes they wear. Nevertheless, the distinction is important as a way of sensitising us to some of the variations in sources of evidence.

Proximate access involves the use of one or more *channels* of access and the adoption of a particular *interactional stance*. The channel or means of observation may be aural or visual, involving listening or looking. The interaction between observer and observed may be 'reactive' or 'non-reactive', depending on the extent to which the observer intervenes or responds in order to communicate with those under observation. In non-reactive research, the observer strives to remain unnoticed as an observer in order to minimise his or her influence on the 'natural' course of events.

[Figure 1.1](#) *Sources of evidence: proximate access*

|                             |                     | <i>Channel</i> |               |
|-----------------------------|---------------------|----------------|---------------|
|                             |                     | <i>Aural</i>   | <i>Visual</i> |
| <b>Interactional stance</b> | <b>Non-reactive</b> | 1              | 2             |
|                             | <b>Reactive</b>     | 3              | 4             |

*Note.* I ignore tactile channels as in social research they are generally adjuncts of aural and visual channels.

[Figure 1.1](#) shows a cross-classification of these two dimensions of proximate access to generate four sources of evidence. Type 1 is exemplified by natural, everyday conversation, in which our sense of hearing is the primary channel through which we learn about others.<sup>2</sup> Non-reactive interaction through a primarily visual channel, type 2, provides evidence of non-verbal behaviour, such as

deportment and manner, and, like type 1, is a form of covert observation. This is not to say that non-reactive researchers habitually hide their physical presence behind screens or one-way mirrors: it is the researcher's presence *as a researcher* which is masked. Similarly, a covert participant observer is, to all intents and purposes, 'merely' a participant in the normal course of events.

Reactive observation through an aural channel, type 3, occurs when the observer questions those under observation in order to elicit responses relevant to the research, and the clearest example of evidence produced in this way would be interview responses. It should perhaps be emphasised that the contemporaneity of observer and observed does not mean that information may only be elicited about the present: the techniques of 'oral history'<sup>3</sup> have exploited interview techniques to obtain evidence about the past. Reactive observation through the visual channel, type 4, involves an attempt to elicit written responses, for example, from those under observation; the observer reads the written evidence in order to acquire knowledge. The clearest example of such research would be the use of self-completion questionnaires, where the subject of the research is invited to answer a series of written questions for later processing by the investigator.

Proximate access therefore characterises a number of the forms of evidence obtained by sociologists through such methods as participant observation, interviewing, and questioning. The communication between observer and observed in situations of proximate access is always direct, even if the observer records this communication for ease of handling. In mediate access, however, the evidence has already become 'fixed' in some material form which the observer has to 'read'. The researcher has no direct access to the situation in which the evidence was produced. When the results of past actions become fixed in a material form



which is capable of survival for a period of time, it is possible for a researcher to study the material medium to obtain indirect evidence about those actions. The central question in considering mediate access, therefore, is the nature of the medium in which the message is fixed. The media may range from solid and substantial forms, such as houses, clay tablets and dead bodies, through the less substantial, such as paper, to completely insubstantial electronic media, which carry their information only for so long as they are supplied with a suitable supply of energy.<sup>4</sup>

Particularly important forms of such evidence have been those used in archaeological studies, where bones and other remains, artefacts, and waste products have been subjected to investigation. The archaeological method typically consists of excavation to disclose the strata of earth and remains which have accumulated on site, successive layers being removed and their details recorded. Items of interest might include the nature of the earth material, holes made by posts, remnants of walls, human remains, and pots, clothing, jewellery, weapons and coinage. Excavation may extend over a large area in order to produce not merely a cross-section but a three-dimensional view which combines vertical stratification with the lateral arrangement of each layer - showing, for example, house plans and street patterns. Through such means as radio-carbon dating, tree-ring dating and fluorine testing of the source material, the archaeologist attempts to construct a relative and, so far as is possible, an absolute chronology of the site. The physical evidence yielded by human and animal remains shades over into that provided by the physical conditions of landform and geology. Historians have shown that the Anglo-Saxon field patterns discovered and used as evidence of settlement and property distribution may be conditioned by an enabling geology.<sup>5</sup>

The use of physical evidence is not limited of course to the distant past. Webb and his colleagues have argued that physical traces may be used for studies of contemporary social action. The number of different fingerprints on the page of a book or magazine, for example, can be used to measure the number of readers, and the contents of dustbins can be used as indicators of consumption patterns.<sup>6</sup>

I have so far considered material traces as indicators of social actions and relations without reference to any intended messages they may contain. Much archaeological research, for example, sees remains and artefacts as unintentional testimony to past actions. When the material media also contain intentional messages, then they may be considered as 'documents'. A document is an artefact which has as its central feature an inscribed text. The material base for the script is irrelevant: it may be a clay tablet, a sheet of paper, or a visual display unit. The term script should not be taken to include only pictorial or patterned decoration - no matter how important these may be as sources of evidence. Script is the written expression of a spoken language and therefore contains a 'text'. The text is the central and most obvious feature of a document, and a book may be considered a paradigm of a document. Inscribed objects such as coins, clocks and cars are not documents in this strict sense because their inscriptions are peripheral to their main significance. There are of course marginal cases, such as stamps, cheques, tickets and gravestones; however, the distinctiveness of documentary evidence should be clear.<sup>7</sup>

Through the various means of proximate and mediate access, the researcher is able to obtain evidence about agents and structures. Observations of natural objects, artefacts (both documentary and non-documentary) and actual behaviour yield evidence about other people. These

three types of source material are intrinsically interdependent: artefacts are natural objects which have been transformed in some way by human behaviour. It is for this reason that many source materials cannot be allocated unambiguously to one type or another, and also, therefore, why research will often involve a combination of proximate and mediate access. For example, a hill or a valley, a natural object, will generally be observed only in the state that results from transformation through human settlement and farming, as sharing a characteristic of artefacts. Even in cases where such obvious signs of human action as fields, footpaths and settlements are not apparent, a landscape may still be partly the result of human activity - as with the North American dustbowl and the Norfolk Broads. Nevertheless, natural geology, geomorphology, and climate will impose constraints on the transformative power of human action, and so any landscape will reflect both natural and human processes.

Much of the evidence available to the social researcher is, of course, analogous to that available to people in their everyday life: people read documents and talk to one another all the time; the issuing of questionnaires is a regular practice in many spheres of life; and police are not the only people to carry out interviews and engage in covert observation. What distinguishes the stance of the social researcher from that of people in their everyday activities is that sociological and historical data are constructed with a scientific, theoretically informed intent; great care is taken - or should be taken - about the quality of the evidence and therefore about the validity and reliability of the data constructed from the evidence. The foundation of scientific research is the quality of the evidence available for analysis. The assessment of its quality is central to the whole argument of this book, and I wish to argue that a simple set

of criteria can be used for this purpose, regardless of the type of evidence. There are four criteria:

1 *Authenticity*. Is the evidence genuine and of unquestionable origin?

2 *Credibility*. Is the evidence free from error and distortion?

3 *Representativeness*. Is the evidence typical of its kind, and, if not, is the extent of its untypicality known?

4 *Meaning*. Is the evidence clear and comprehensible?<sup>8</sup>

*Authenticity* is a fundamental criterion in social research, as questionable sources of evidence can mislead the researcher unless he or she knows that they are inauthentic. The field-worker carrying out interviews, for example, must be sure of the identity of the interviewee before undertaking the interview, and the participant observer must know that he or she is in the right locale for the observations to be undertaken. Similarly, users of the physical evidence of human remains must be sure of the authenticity of those remains – Is the skull that has been discovered really one of a hominid, or is it, perhaps, an animal skull or a plausible forgery?<sup>9</sup> The secondary analyst handling questionnaires from an earlier study will want to know whether the material available is actually from the study in question (its ‘known provenance’), where it has been stored since the original study, and whether there has been any editing or copying from the original. Unless the researcher is able to come to a conclusion about the authenticity of the evidence, there is no possibility of an informed judgement about the quality of the data eventually constructed.

*Credibility* refers to the extent to which the evidence is undistorted and sincere, free from error and evasion. Interviewers must be sure that their interviewees are taking the interview seriously and are therefore saying things which can be regarded as *prima facie* credible. If the

interviewee is lying, or regards the interview as a humorous diversion, the responses must be handled differently from those acquired in a serious and sincere interview. Participant observers must know whether the things they are observing are credible events for that situation or are, perhaps, staged for their benefit and are therefore to be regarded in a different light. The secondary analyst will want to know whether responses to questions on, say, sexual behaviour which were asked thirty years ago are more or less likely to exhibit coyness and oblique answers than responses to similar questions asked today.

*Representativeness* refers to the general problem of assessing the typicality, or otherwise, of evidence. But it should not be assumed that the researcher always desires 'typical' evidence: what is important is that the scientist should know how typical it is in order to be able to assign limits to the application of any conclusions drawn.<sup>10</sup> The researcher handling physical evidence, for example, will want to know whether the surviving artefacts on the site (such as jewels and tools) are typical of the range of artefacts once in use by those living on the site. Have clothes or tools made of more perishable materials failed to survive the years buried under the ground? A central issue is the nature of any sampling techniques used by the present researcher or by those responsible for granting or securing access to the sources. If interviews are to be carried out on a random sample of workers, then the researcher must be confident that the sampling frame and procedures are adequate for the production of such a sample; and the secondary analyst of questionnaires will have to know what sampling methods were employed in the original study and whether the archivists or clerks responsible for storing the schedules have used additional selection or sampling methods to reduce the bulk of the material to be stored. In the latter case it is important to

know whether any schedules have been destroyed and, if so, whether this was because of storage problems or in order to hide some flaw in the initial research design.

The final quality control criterion, *meaning*, refers to the extent to which the evidence is clear and comprehensible to the researcher: what is it, and what does it tell us? The observer of arcane tribal practices in the Sudan, for example, will need to ask such questions as 'Is this witchcraft?', a matter of how the actions under investigation are to be described.<sup>11</sup> The contemporary observer of violence at football matches will similarly need to know whether the 'hooligans' involved are engaged in 'ritualised' aggression or 'real' violence. An observer may be sure that he or she is in the right place, is not being faced with staged events, and is not observing unusual events, but may still be unclear as to *what* those events are and how they are to be described. In handling physical evidence, the researcher will ask such questions as whether the position of a body indicates ritual killing or domestic violence, and whether the pattern of holes discovered on a site should be seen as the remains of a dwelling or a temple. Finally, the secondary analyst can only assess the evidence available from the schedules if the original instructions given to interviewers and coders are known; without this knowledge it may be impossible to interpret the meanings of particular responses.

Data construction involves the use of methods and procedures for the derivation of data from evidence of known quality. In the natural sciences it may often involve systematic theories of observation and measurement, but this is far less common in the social sciences. In the absence of theories of measurement, resort must be made to rule-of-thumb techniques based on implicit, everyday theories. Accepted theories of measurement are far from evenly developed in the social sciences. While the

assessment of authenticity and the construction of authentic data through interviewing and the use of documents often involves implicit and loosely established techniques, the archaeological use of physical evidence involves rigorous theories of authentication rooted in accepted scientific knowledge about, for example, the behaviour of carbon and fluorine atoms. Areas of social research where well-founded theories of measurement and data construction have emerged are those concerned with the assessment of representativeness and, to a lesser extent, credibility and meaning.<sup>12</sup>

The theory of sampling, for example, is well established. It was introduced into the social sciences by A. L. Bowley,<sup>13</sup> a statistician who carried out a follow-up to poverty studies undertaken by Booth and Rowntree.<sup>14</sup> While the earlier writers had each attempted to survey the whole population of their area, Bowley appreciated that advances in the theory of probability showed that a randomly drawn sample from the population could yield valid information about the population as a whole, even when the sample size was fairly small. Sampling theory, therefore, is an attempt to spell out the grounds on which generalisations from the sample to the wider population can legitimately be made. Bowley's intention in introducing sampling theory, apart from reducing the time and cost necessary for social research, was to make it possible to calculate the size of sample necessary in order to achieve a satisfactory degree of accuracy.<sup>15</sup> The researcher can be confident of producing accurate results so long as he or she is confident that the sample is unbiased. 'Bias' depends upon the representativeness of the sampling frame, the randomness of the method of sampling, and the number of sampled cases which are unavailable, inaccessible, or fail to respond. Because numerical values cannot be assigned to these



factors, the assessment of bias is always a qualitative judgement, but it is nevertheless based on an understanding of the theory of sampling.

There is no single, widely accepted theory for the measurement of meaning. There are, instead, a number of competing and complementary theories alongside a number of more generally accepted principles. 'Measurement' should be here understood to refer to the processes of coding and classifying source material into the theoretically defined categories required for the researcher's purposes. Measurement is not always a quantitative procedure. The absence of generally accepted procedures for the measurement of meaning has led many researchers to treat coding as if it were a theoretically neutral process in which the sole consideration is the convenience and parsimony of the categories employed. Researchers have often taken over administrative categories, such as those for criminal offences, and have failed to consider the correspondence, or lack of correspondence, between these administrative categories and the sociological concepts in which they are interested. To assume that this relationship is unproblematic, and that coding categories are mere technical devices, is to ignore the theoretically grounded nature of scientific research and to run the risk of what Mills has termed 'abstracted empiricism'.<sup>16</sup> It is, of course, the case that a useful coding system must meet certain technical criteria - the categories must, for example, be exhaustive and mutually exclusive - but the codings are essentially ways of operationalising theoretical concepts. Research cannot be theoretically and empirically progressive unless the concepts of the theoretical framework are adequately operationalised. It is for this reason that questions of coding rest upon theories of meaning which tell the researcher how to 'read' and interpret evidence.<sup>17</sup>

# What are documents?

Questions on the nature and use of documents have figured prominently in the methodological writings of historians, but it would be wrong to see documents as exclusively historical sources. Historians have often been concerned with only a very narrow range of documents, especially State documents of constitutional or diplomatic origin. Nevertheless, the methodological views of the classic historians remain the essential starting-point for a more general consideration of the use of documentary sources in social research.

In their classic compendium of research practice, Langlois and Seignobos remark that 'Documents are the traces which have been left by the thoughts and actions of men (*sic*) of former times',<sup>18</sup> and it is only through these traces, they argue, that we can know the past. 'For there is no substitute for documents: no documents, no history'.<sup>19</sup> This contention is rooted in the nineteenth-century revolution in historical writing initiated by Ranke, and it became the corner-stone of professional academic history. Written documents discovered in libraries and archives were regarded by Ranke as superior to observational and archaeological evidence, and to reminiscences and oral traditions. Learning to handle documentary evidence became the central feature of the research training of the historian because it was seen as the characteristic method of history, the method which distinguished it from other disciplines and gave the voice of authority to the expertise of professional historians. The historian was an expert in the handling of documents, and this expertise was a bulwark against the intrusion of uninformed judgements from outside the discipline.<sup>20</sup>

The rationale for preferring documentary evidence to all other sources of evidence about the past is, in part, a sheer matter of survival. For the distant past it is generally the case that interviewing and observation are impossible,<sup>21</sup> and that documents have survived in great numbers. The expansion of the State and the growth of the economy in the modern period have generated massive quantities of written material which appear to give a direct insight into past events. These documentary survivals are regarded as especially valuable because they 'are not deliberately designed for the benefit of the historian'<sup>22</sup> and so can be seen as the objective residue of the past; they provide the historian with the unwitting testimony of people in the past. It has, nevertheless, been recognised that such documents do not speak for themselves. The task of the historian is to speak objectively on their behalf. Langlois and Seignobos claim that the craft of the historian lies in moving from the *source* to the *fact*: 'The document is his starting point, the fact his goal.'<sup>23</sup> These writers voiced the prevailing view in German historical writing of the late nineteenth and early twentieth centuries, when the historian's task was seen as that of hermeneutic reconstruction. Echoing Dilthey's philosophical underpinning of history, they argue that the historian must 'revive in imagination the whole of that series of acts performed by the author of the document'.<sup>24</sup> Only by doing this can the document be placed in context and an understanding of its meaning and significance be achieved. While many of their contemporaries might not have accepted the philosophical notion of 'reliving' the experience of others, the general point was accepted that the historian had to place the document in the context of its conditions of production before an appraisal of its message could be made.

As has already been remarked, the range of documents considered by historians of the classical school tended to be remarkably narrow. The exemplary documentary source might be the manuscript report of a minister to a king, with the reports, letters and diaries of those active in the service of the State coming a close second. Printed documents of all kinds were poorly regarded, though Acts of Parliament and other official records might be used in studying the more recent period when the printing of documents had become more routine. But only those documents concerned with constitutional and diplomatic events were highly regarded. An important feature of the development of the discipline of history in more recent years has been a broadening of attention to other documentary sources, allied with the growth of economic and social history, and a willingness to supplement documents with other sources of evidence.

The legacy of the classical tradition has had a great influence, however, on attempts to define what exactly a 'document' is. A particularly influential view among those social researchers who have considered such matters is that of Sidney and Beatrice Webb, who argued that the manuscript and printed sources which the social researcher may legitimately consult can be divided into two classes: 'documents' and 'contemporary literature'. A document, they argue, is 'an instrument in language which has, as its origin, and for its deliberate and express purpose, to become the basis of, or to assist, the activities of an individual, an organisation, or a community'.<sup>25</sup> Documents are, therefore, 'exclusively for the purpose of action' and are not written to inform historians, sociologists or any other detached observer. The Webbs are here generalising the classical historian's view from the realm of the State to all organised groups and all purposive action. 'Documents' are the accounts, returns, statutes and proclamations that individuals and groups produce in the course of their

everyday practice and that are geared exclusively to their immediate practical needs.

'Contemporary literature', on the other hand, is a residual term for all other written sources, such as treatises, sermons, newspapers, poems and biographies, which are contemporary with the events or people under investigation. The Webbs argue for the primacy of 'documents' over 'literature' in the study of social institutions on the grounds that it is never possible to be sure that the literature accurately and unbiasedly records what actually happened or what is contained in the original documents. 'Literature' does give, however, the 'background' which is often missing from the documents, and so enables the researcher to uncover, for example, the conflicts and struggles which lay behind the apparent consensus of a government statement.

There are clearly a number of problematic features in the Webbs' argument. While they correctly recognise the possible sources of inaccuracy and bias in contemporary literature, they fail to see that similar problems beset the 'documents' and that they cannot, therefore, be given the privileged status conventionally accorded to them. An adequate approach to the use of documentary sources must adopt a more general definition of 'document' than that of the Webbs and the classical historians.

The position that I wish to take here is that a document in its most general sense is a written text, as defined in the previous section. Writing is the making of symbols representing words, and involves the use of a pen, pencil, printing machine or other tool for inscribing the message on paper, parchment or some other material medium. The introduction of paper in place of clay, stone and parchment as the receptacle of writing, and the invention of printing as a supplement to handwriting created the archetypal document: the text printed or handwritten on paper. But to

recognise this as the archetype or exemplar does not mean that it is not possible to regard hieroglyphic and cuneiform tablets, for example, as documents in essentially the same sense. All that differs in these cases is the physical embodiment of the document. Similarly, the invention of magnetic and electronic means of storing and displaying text should encourage us to regard the 'files' and 'documents' contained in computers and word processors as true documents. From this point of view, therefore, documents may be regarded as physically embodied texts, where the containment of the text is the primary purpose of the physical medium.

This purpose, however, is the purpose of the author of the text, and it is always necessary to establish through empirical research what purpose may, or may not, lie behind the production of artefacts. As the imputation of purpose may always be thrown into question, it is inevitable that the borderlines of the category 'document' will be somewhat fuzzy. The inclusion of an inscription on a coin, for example, does not necessarily make the coin a document, as the primary purpose of the inscription may simply be to denote the monetary value of the coin in transactions and to identify its country of origin. Even so, coins sometimes contain especially fulsome inscriptions which may have been intended, for example, to proclaim the authority and majesty of an emperor. Similarly, gravestones have the primary purpose of marking a place of burial though are sometimes inscribed with texts which serve as documents.[26](#)

This hazy borderline surrounding the archetypal document is especially obvious in the case of so-called printed ephemera. Advertisements, handbills, invoices and so on are clearly recognisable as documents, but the status of bus tickets, stamps and postcards is less clear-cut. The latter are circulatory or commercial devices, like coins, but are

embodied in paper and so bear a closer physical resemblance to the archetypal document. A particular problem arises with maps, paintings, films and similar sources. Maps, for example, embody a text in pictorial and written form and are indistinguishable, in this respect, from other documents. Oil paintings on the other hand might be regarded as aesthetic remains rather than documents, though this is not to deny their importance to the social researcher. Photographs fall on the borderline between these two; whether they are the aesthetic products of the studio photographer or the physical residue of ritualised holiday snapshots, they are often used by families and organisations as documentary records of events to be stored in an album or archive. My aim in this book is to recognise this diversity in documentary sources as a valuable feature of social research. The discussion in it will concentrate on the mainstream of written documents, but the general principles of documentary research apply equally to those on the borderlines.

It is possible to classify documents by their content – as business, political, religious, etc. – but this tends to result in a myriad of overlapping categories, as many documents contain information on more than one area. Therefore, it is preferable to explore the range with a more analytical approach. [figure 1.2](#) uses the two dimensions of *authorship* and *access* to generate a typology of modern documents.<sup>[27](#)</sup> The dimension of authorship refers to the origin of the documents, and its applicability is clearly dependent upon the existence of a separation between the ‘personal’ and the ‘public’ or ‘official’ spheres and, within the latter, between the State and private bureaucracies. The distinction between the personal and the official world was central to Weber’s discussion of the rise of modern bureaucratic administration, with its separation of the ‘household’ from the ‘office’.<sup>[28](#)</sup> This differentiation is a



hallmark of the modern period and the classification in [figure 1.2](#) has less relevance to the medieval and earlier periods, when the distinction between ‘personal’ papers, ‘Church’ papers, and ‘State’ papers was not made at all sharply – if, indeed, it was made at all; and it is only in recent history that official prime ministerial documents in Britain, for example, have ceased to be regarded as the personal property of the Prime Minister.<sup>29</sup>

[Figure 1.2](#) *A classification of documents*

|               |                | <i>Authorship</i> |                 |              |
|---------------|----------------|-------------------|-----------------|--------------|
|               |                | <i>Personal</i>   | <i>Official</i> |              |
|               |                |                   | <i>Private</i>  | <i>State</i> |
| <i>Access</i> | Closed         | 1                 | 5               | 9            |
|               | Restricted     | 2                 | 6               | 10           |
|               | Open–archival  | 3                 | 7               | 11           |
|               | Open–published | 4                 | 8               | 12           |

The dimension of ‘access’ refers to the availability of the documents to people other than their authors. Documents subject to ‘closed’ access are those which are available only to a limited circle of eligible insiders, normally to those who produce them and to their bureaucratic superiors. By contrast, ‘restricted’ documents are accessible on an *ad hoc* basis under specified conditions to those outsiders who are able to secure the permission of insiders; they are therefore, normally closed to outside access, though their authors or custodians may be willing to grant access on application. ‘Archival’ access exists where the documents have been lodged in a place of storage which is open to all comers; researchers and the general public therefore may use an archive subject only to minimal administrative restrictions – such as the need to apply for a reader’s ticket, supply