

“A” Level Sociology

A Resource-Based Learning Approach

Module One: Theory and Methods

Unit M5: Secondary Methods

Introduction

Unit M5

In the two previous Units we looked at **primary methods** of research and, to complete the **Methods** Units of this Module we need to now look at **secondary methods of research**.

- To begin with, we can broadly define **secondary data** as any information that has **not** been generated personally (first-hand) by the researcher who uses it.
- A **secondary method** of research, therefore, is one used by a researcher to **gather data that already exists** in some shape or form. For the sake of clarity, this Unit focuses on **secondary sources** of data rather than **secondary methods** of data collection.

All this means, in effect, is that we will be focusing on the various ways sociologists can **use** data that already exists. This is because, **logically**, all data - sociological or otherwise - is ultimately produced by a **primary method** (that is, someone, somewhere and at some time had to create the data).

- Sociologists use secondary sources for a variety of reasons, as we will discuss in more detail in a moment, and the **sources** of such data are many and varied.
- These range from highly **quantified** sources such as **statistics** (both **official** and **unofficial** - see below) to the more **qualitative** such as **personal documents** and **diaries**, **government** and **business reports** and the **mass media**).

The Mass Media, for example, involves a huge range of potential sources - **newspaper** and **magazine articles**, **television news**, **reports** and **documentaries**, **films**, **plays**, **novels** and the like

In addition, secondary sources can be both **contemporary** (that is, they belong to the present) or, most-usefully, **historical**.

- It's not particularly difficult to see why sociologists use secondary sources and we can briefly note a couple of reasons for their use:
- Firstly, on a purely **practical level** the use of secondary sources may represent a substantial saving of **time**, **money** and **effort**.

The basic argument here, therefore, is that it may be unnecessary or impractical for a researcher to create some forms of data using primary methods when such data already exists

For example, in Britain the government collects a huge amount of statistical data about national and local trends in things like employment, crime, education and so forth and distributes it freely in published documents. Thus, a researcher, for the price of a book or a visit to a public library has immediate access to statistical data on patterns of crime, for example, that would cost an enormous amount of money time and effort to collect personally.

- Secondly, on a **methodological level** the use of secondary sources may be a necessity if **historical** and / or **comparative** research is being carried-out.

- **For example**, the following are two examples of research that have used, in the first instance, **historical** evidence to support an argument about the nature of **childhood** and, in the second, **comparative** data to support an argument about the nature of **suicide**...

Classic Study:
“Centuries of Childhood”: Philip Aries, 1962.

To test his theory that “childhood” was a relatively recent phenomenon, Aries used a variety of historical accounts and sources to demonstrate that, in the past, “childhood” (as we understand in modern societies) did not exist. Aries’ research drew upon such diverse sources as paintings, contemporary documents and the like and secondary source materials

Classic Study:
“Suicide: A Study in Sociology”: Emile Durkheim, 1897.

Durkheim was interested in the theory that suicide had social, as opposed to psychological or pathological, causes and to test his ideas he collected statistical data from a range of different countries. Without the existence of such data Durkheim would not have been able to carry-out a comparative study of suicide rates and causes. In addition, of course, given his subject matter, without using secondary sources Durkheim would have been unable to effectively research suicide since primary data in this area does not exist.

Although there are clear and justifiable reasons for using secondary sources of data, we also need to be aware of possible **drawbacks** and **limitations**.

We need to note, for example, that, as a general rule-of-thumb, whenever we use **any** data we need to consider a number of important **methodological concepts** that surround the collection and use of such data. For example, whenever we collect data we should always consider the concepts of:

- Data **reliability**:

As a general rule, **quantitative** data tends to be considered as more **reliable** than **qualitative** data.

- Data **validity**.

In relation to **validity**, the opposite tends to be true. **Quantitative** data, for example, may be **too limited**, narrow and restricted in its scope to allow us to capture the full richness and depth of social interaction.

If we study a similar group we cannot be sure that any **differences** between the two groups are so great and **uncontrollable** that we are comparing “**like with like**”. This is because **qualitative** data, by its very nature, is more difficult for sociologists to **replicate**. A **covert participant observation** study, for example, **cannot** be **repeated** in exactly the same way as the research was originally carried out for fairly obvious reasons: the group being studied will change over time or it may not exist anymore.

A **questionnaire or structured interview**, for example, may so **limit the responses** that can be given by a **respondent** that they fail to present us with a true picture of the reality of someone’s life, thoughts and opinions.

Statistical data, on the other hand, may provide us with a **snapshot** of some aspect of human behaviour as it was at the moment the statistics were collected - but they cannot tell us anything much about why people behave in particular ways

In addition we should also consider what **Scott** ("**Documents in Social Research**": **Social Studies Review**, 1990) has identified as **crucial concepts** in the **evaluation** of secondary sources.

1. Authenticity

In simple terms, we need to be certain of the **source** of the data we use. We need, for example, to be sure that a document is not a **forgery**. In addition, we need to know whether we are working from **original documents** or from **copies** of such documents (which may have been changed, subtracted from or added to by other authors).

2. Credibility

For example, we may want to know **who** wrote something and why they wrote it (see **purpose**). We would need to know whether or not the author had **first-hand experience** of the things they describe or whether they were simply repeating something "**second or third hand**" ("hearsay" evidence). We would also have to consider the **representativeness** of the source in this context.

3. Representativeness

When considering a secondary source we need to know, for example, if it is simply **one individual's view** or whether it is **representative** of a whole **range of views**. **Newspaper** articles, for example, may simply be the **personal, unsupported and unrepresentative**, view of a **single journalist**. Similarly, **historical documents** tend to reflect the views of **particular social classes, religions** and so forth (mainly because it was the upper classes and the religious who recorded their views). Finally, perhaps, a single document which may be the only surviving contemporary record of an event will provide us with an insight into that event, but without **supporting evidence** it is unlikely to give us a representative view of the event.

4a. Meaning (1)

To use secondary sources in sociological research we invariably have to consider their **meaning**. This can be done on two main levels

Firstly, we can consider their **literal meaning** (that is, what they actually say about something). Secondly, however, secondary sources can have a **metaphorical meaning** in the sense that they can be used to **illustrate a particular sociological point**. In this sense, data does not always have to be taken at face value (what it actually says). **Historical documents**, for example, can be interesting to sociologists more for what they tell us about the things that concerned people in the past than for their actual content.

4b. Meaning (2)

In addition to the first meaning, we can also use documents **comparatively**; that is, we can compare different views (for **example**, between the **past** and the **present**) to illustrate the sociological point that people tend to view the past nostalgically and they use this

nostalgia to contrast an **idealised** view of “the **past**” with a **present** that is **dangerous and uncertain**. Consider, for example, the following two accounts of family life:

"Family life is collapsing and responsible parents can no longer afford children", the government was warned yesterday. And lack of parental control and guidance lies behind many of today's pressing social problems, said Lord Joseph, the former Education Secretary, Sir Keith Joseph.

"Part of the background to crime, to drug addiction, to low motivation at school, to poor job prospects and to the transmission of all these problems to the next generation comes from inadequate parenting. If you want to destroy a country, you debauch its currency - and the way to destroy a society is to destroy its children."

M.Benns "Save our Children from the Collapse of Family Life": **Daily Mail 29/06/90**.

"The withdrawal of women's attention from the care of her offspring, and from domestic duties is an unnatural arrangement and a stigma upon the social state. Young children are left at home under very inadequate conduct and almost without restraint. They are left to play at will and to expand into every lawless form. Ignorant of cooking and needlework, unacquainted with all that is necessary to promote the comfort and welfare of a home... slatternly and ignorant. The unfortunate man who marries a woman of this class suffers also. There is neither order nor comfort in the home and his meals are so irregular and ill-prepared and his own fireside presents so few attractions that he is tempted to the beerhouse. The social evils are aggravated by the independence of the young of both sexes. The child receives his wages on his own account and in some cases he will even remove from the parent roof. It is palpably a system fraught with innumerable evils, especially when we consider the early direction of the child's mind to the value of money and the consequent temptation to procure it by illicit means."

T.Beggs "An Inquirey into the Extent and Causes of Juvenile Depravity", **1849**

5. Purpose

Whenever we use secondary sources, (unless it's the work of other sociologists), it's highly unlikely the data will have been produced with sociologists in mind. In this respect, the producers of the data will have had their own reasons, concerns, interests and agendas for producing it and we must be aware that these may not coincide exactly with sociological concerns, interests and agendas. Official employment statistics, for example, are produced by governments who may want to show (for political reasons) that there is very unemployment in our society. The statistics they produce may reflect this concern and they may not, therefore, give us a valid, sociological, picture of “true levels” of unemployment.

- Although, for the purpose of illustration, I've outlined these ideas **separately** (as if they were **discrete** or **self-contained**), it is clear that, in reality, we will need to consider how each **connects to** (and has an **effect on**) the others when we **evaluate** secondary sources of data.

For example, when we consider the **authenticity** of a data source we will necessarily have to consider its **credibility** as a source, how **representative** it is and the **purpose** for which it was originally produced.

- When we carry-out **primary research** most - if not all - of the above ideas will be considered as part of the research process. By and large, therefore, we have, as researchers, **control** over these things (whether or not we choose to completely follow them is, of course, another matter).
- When dealing with **secondary sources**, however, it is not always so easy to ensure that the data we use is **reliable** or **valid**, **authentic** or **representative**.

For this reason, we have to be especially **cautious** about the use of secondary data in our research.

Having noted these important **methodological concepts** we can now start to examine **secondary sources** of data in more detail. As with the previous Units, we can most easily and conveniently organise such methods in terms of **quantitative** and **qualitative** categories.

Thus, we can begin by looking generally at **secondary quantitative sources** of data and specifically by looking at **statistical data**. In this respect, although there are many potential sources of **statistical data** that can be used by sociological researchers (for example, **governments, social scientists, journalists, media organisations, businesses** and the like all, at various times, **produce statistical data** that can be used, in different ways, as sources of secondary statistical data), we can, for the sake of convenience and illustration, generally consider the most important sources of statistical data for the sociological researcher to be those produced by "**official sources**".

By **official**, in this context, we mean statistical data produced by **governments**. For this reason, we frequently see such secondary data referred to as "**Official Statistics**". You should **note**, however, that the points being made in this section relating to official statistics can, with a little care and thought, be applied to any statistical data

In **Britain**, for example, the two main sources of official statistical data we have is that produced by:

- **Government departments** (such as the Home Office and the Department for Education)
- **Government agencies** (such as the police),

By and large, this **demographic data** (for example, data about such things as where and how different individuals and groups live in a society) is collected by the government to be used as the basis for the creation of government policies.

By understanding something about the way people live the government is able to adjust both its **policies** and its levels of **expenditure** to take account of **changes** in society.

For example, whether the population is generally increasing - and by how much - whether there are an increasing number of elderly people in the population and so forth.

This statistical data is published annually by the government in "**Social Trends**" (HMSO). Data about areas such as **crime, work and leisure, family life, education, politics, religion** and so forth is routinely published in Social Trends, which makes it an invaluable **source of secondary data** for sociologists.

In addition, "**The Annual Abstract of Statistics**" contains a wealth of statistical data compiled and presented on an **historical** basis (for example, it may allow us to **compare** crime statistics over a thirty year period), while "**Regional Trends**", as the name suggests, provides **statistical** data on a **regional** basis.

As with any method of data collection used by sociologists, care must be taken over the **use** and **interpretation** of statistics, whether they be from "**official**" sources (such as Government departments) or "**unofficial**" sources (such as the work of social scientists, journalists, business organisations and so forth).

Having noted this, we can start this section by looking at why sociologists use official statistics.

Bilton et al ("Introductory Sociology") suggests five main reasons for the **use** of official statistics as a source of data for sociologists:

a. Availability.

It may be the case that official statistics are the only available source in a particular sociological area of interest. **Emil Durkheim**, for example, in his classic study of suicide ("**Suicide: A Study In Sociology**", 1897) used official statistics drawn from coroners' reports from a number of different societies in the attempt to establish that suicide rates varied within and between societies. By so doing, he was able to argue that social factors, such as religious belief, were significant variables in the explanation of why people committed suicide.

b. Readily available.

The researcher does not have to spend time and money collecting his / her own information.

This may be an important **practical research consideration**.

c. Examination of trends / changes over time.

Using statistical data drawn from a number of different years it is possible to see how something has changed over a long period. For example, statistics of educational achievement can show us changes in relative levels of academic achievement between boys and girls.

Again, this may be an important practical research consideration if the sociologist is concerned with attempting to analyse variations in behaviour over a long period of time.

d. Comparisons.

Statistics can be used for Inter-group comparisons (for example, the examination of differences in middle-class and working-class family size), as well as cross-cultural comparisons (for example, a comparative study of crime rates in different countries).

This kind of information may well be too expensive and time-consuming for the sociologist to collect personally, for example.

e. "Before" and "after" studies.

For example, we could use statistical data to examine the effect that changes in the law have on patterns of divorce by noting the number of divorces **before** the legal change and doing the same for divorces after the change.

- As the above suggests, there are a number of very good reasons why sociologists should use official statistics and these mainly relate to various aspects of **practical research considerations** when choosing a method of data collection to use. However, the **uncritical** use of official statistics may involve a number of **problems**, not the least of these being that official statistics are invariably collected by non-sociologists (the significance of this being that they are collected for a specific purpose - to provide governments with information - not for the convenience of sociologists who may well use different categories to those used by statisticians).

In this respect, as with any source of data, we have to be aware, that Official Statistics are both a valuable source of information (if used **critically**) and a potential source of biased data (if used **uncritically**).

What we need to do next, therefore, is to look at a number of potential limitations (both **practical** and **methodological**) involved in the collection and use of official statistics. These fall into four main categories:

1. Definitions.

In relation to the way in which official statistical categories are defined, we have to be careful in two main respects:

- a. The definitions used by the collector of official statistics may not be the same as those used by the sociologist.

For example, we need to ensure that the official definition of such concepts as "crime", "unemployment" and "class" is the same as our own.

- b. The basis for the collection of statistics by governments may change over time.

This is significant if your research is concerned with, for example, comparisons of employment / unemployment levels between the present and the past. Between 1980 and 1990, for example, the British government changed the way it defined unemployment (and hence the way it which it collects official data about the unemployed) approximately 25 times...

2. The Purpose of Official Statistics

Although **official statistics** are collected by **governments** in order to provide information about what's happening in society, it would be **naïve** to pretend that the collection of such data is **unaffected** by **political and economic considerations**.

As in the example above, the British government has changed the way in which it defines unemployment a large number of times over the past few years. It may be a happy coincidence that all but one of these changes (the first) has resulted in a fall in the number of people defined as being "unemployed", but, there again, it might not...

However, I'm not suggesting that the data provided by Official Statistics on unemployment is collected in a statistically-faulty way. That is, there is no suggestion that the data is inaccurately collected, as such (that is, that the **statisticians** who are charged with collecting such data deliberately set-out to mislead). Rather, it is to suggest that the way in which one defines a "problem" will affect the way in which one collects data about it - this is a (sociological) problem of **methodology**.

Thus, whilst we may have a valid picture of such things as levels of car theft and murder, the picture that we gain from the statistics about such things as levels of vandalism, burglary and so forth may not be valid...

3. A partial picture of reality?

When using statistics of any kind, we need to be aware of two points:

a. Firstly, that any statistical account will represent a "**snapshot**" of social interaction as it was at the moment the statistics were compiled. Just as a photograph captures a single frame of social activity, so too do statistics capture a fragment of any interaction process. This may be significant if we are using statistical data to track developments in a social process.

b. Secondly, and perhaps more importantly, the statistical data we use may only represent a **partial picture of reality**. An **example** may help to clarify this point:

Data collected by the police, for example, is only likely to represent a proportion of the true level of criminal activity / behaviour in our society, because although the police have a duty to record crimes that are notified to them (hence, as an aside, we can be reasonably certain that the statistics are relatively reliable), not all crimes (for whatever reason) are notified to the police (which casts some doubt upon the validity of official crime statistics).

4. Interpretation

In order to understand any form of statistical analysis, it is evident that we have to interpret their significance - that is, we have to understand the ways in which they are meaningful to both ourselves and others. We can illustrate this idea by looking at the notion of "clear-up" rates within official statistics about crime.

Thus, if we look more closely at crime statistics in order to analyse such things as police "clear-up" rates (that is, the number of reported crimes solved by the police), we have to be careful about the way in which we interpret the data because the official definition of a "solved crime" may not accord with our definition about what constitutes a solved crime...

Officially, therefore, the police definition of a solved crime is one in which an arrest is made. If the arrested person is subsequently found to be innocent, this makes no difference to police clear-up rates (and leads to the statistical anomaly that it is quite possible for the same crime to appear more than once in the clear-up statistics).

In addition, a further problem of interpretation is involved when we consider that such statistics are not always either reliable or valid in terms of the police actually catching criminals.

For example, it is not unusual for the police to "plea bargain" with arrested individuals. When, for example, you read in court reports that,

"Mr. X pleaded guilty to burglary and asked the court to take into account 55 other similar offences...",

it may be indicative of the fact that Mr. X admitted to these "similar crimes" in return for the police charging him with a lesser offence (one that carries a lighter sentence, for example). Mr. X may not actually have committed all these crimes, but the police may work along the lines that they've managed to catch someone

who is an admitted burglar and he probably committed a number of crimes for which he wasn't caught, so in return for the possibility of a lighter sentence Mr. X could be induced to admit to these "unsolved" crimes...

Whatever we may think about the morality of such practices, we have to be aware that they exist and may have an affect upon the way in which official statistics relating to crime are compiled.

- In relation to secondary sources of qualitative data it is evident that these are many and varied, since almost all information that people produce in their everyday lives could be included in this category of data collection. However, in relation to sociological research, we can limit this category to a number of major sources that we can briefly note as follows:

1. Historical Documents.

Sociologists, as you will appreciate as the course develops, are frequently concerned with the analysis of historical events. This is mainly because we tend to see social development as a long, slow, continuous, process. Much of what we experience in our society today, for example, has its origins in the desires, hopes, fears and struggles of our ancestors...

Historical analysis is useful not only for our understanding of how societies develop, however; it is also important as a source of comparison (and comparative data, is an important aspect of most sociological research). Using historical documents, for example, we can understand how people lived in the past - both in our own and other societies (allowing us to make comparisons over both time and space). We can use such research data to understand how society has changed, how it has not changed and, perhaps, how it might change in the future (although this is a rather contentious idea...).

In addition, we rely upon historical documents as a (secondary) source of information simply because it is not possible for us to collect primary data about the past. Having said this, it should be evident that historical documents (like any potential source of data) are liable to be incomplete, inaccurate or simply partial (written to reflect one particular point of view, for example). This is as true of "official" (usually governmental) historical documents as it is of "unofficial" (personal) documents.

This type of documentary evidence is normally of the qualitative type (reports and accounts of people's behaviour and so forth) although, of course, some historical documents may attempt to quantify certain forms of behaviour (the Domesday Book produced for William the Conqueror, for example, represents an attempt to quantify ownership of land in Feudal Britain after 1066).

2. Personal Documents.

While a great deal of historical documentation consists of government reports, enquiries, commissions, surveys, "official statistics" (births, deaths, marriages, divorces, etc.) and parish and Church records, a further source of qualitative data is to be found in such things as personal letters, diaries and books.

We could also note, in this context that personal documents might include such things as oral histories, novels, films and the like.

Again, personal documents are likely to be both one-sided and incomplete (and all tend to relate more to the personal lives and experiences of individuals), but they nevertheless frequently represent a significant - and at times only - source of data for sociologists interested in historical analysis and comparative research.

3. The Mass Media.

Like historical and personal documents, the mass media are a readily available source of secondary data - much of it highly opinionated and deeply subjective...

Analysis of the media (in all its different forms) often tells us more about the thinking and prejudices of various powerful individuals and groups, of course, but it nevertheless represents a significant source of qualitative data.

On a more personal level, the mass media will be very useful to your A-level studies as a source of contemporary data, since it frequently carries reports, analysis and comment on relatively up-to-date social research.

Having outlined some aspects of secondary source material in this set of Notes, further Notes look in more detail at specific types of secondary material (such as official statistics). There are also a number of separate sets of Notes covering areas such as experiments and participant observation.