Disasters, Vol. 3, No. 1, pp. 13-15.
Pergamon Press Ltd., 1979. Printed in Great Britain.

Social Investigation: Some Preliminary Considerations

With an increasing preoccupation with science and scientific method, modern sociologists seek to ground their theoretical constructs in empirical facts. So sociological enquiries have become a prominent aspect of sociology, and are as much related to sociology as technology is to science. Moral philosophers may feel very strongly about the rights and wrongs of applying techniques to society. Thus, although an analysis of social problems may suggest avenues of action: and may, therefore, be a necessary prerequisite to the information of social policy, it cannot claim – as some idealists might – to provide a prescription for action. Indeed analysis (like medical diagnosis) may be one thing and cure quite another. To know what ails society (or a patient) is not saying that one can necessarily cope with the problem. Moreover, the social scientist, perhaps more than the technologist, runs the danger of becoming a social engineer – in other words a dictator who manipulates society to bring about ends desired on theoretical consideration. At present such danger is small because sociology is not comparable to the natural sciences in its accuracy of prediction, or its ability to discover laws of society that operate with the exactitude of the laws of nature. However, if sociology should ever reach such precision some people would claim that this would also be the end of societies as living entities who would be replaced by mechanically planned processes no longer concordant with what we today describe as "human".

And of course it is necessary to be particularly conscious of potential social engineering in disaster situations, where outside agencies are in danger of being too manipulative.

Nevertheless, the detailed study of social problems can provide important insights and is of tremendous value in evaluating the effectiveness of social policy and social planning. To discover, for instance, what health services are achieving. methods of sociological enquiry can be effectively used, and some other areas where sociologists make useful team members in research are: environmental disease; mental health; maternity and child welfare; delivery of health care; occupational health etc.; and planning for economic development.

The basis of such investigations is quanititative assessment and thus involves a fair amount of statistics. But it cannot be stressed too emphatically that on the whole it is far better to consult expert statisticians than to be subject to misinterpretation of data by trying to handle complex problems in simplified manner. The most important asset that any research worker can have is both a healthy respect for and great distrust of the numerical expert. If the statistical analysis runs counter to the scientific expectations and hypothesis, this need not mean that the hypothesis was wrong, but more often than not it is due to faulty techniques of survey design. This paper, therefore, will be concerned to discuss, in detail, questionnaire design and theories of sampling and leave readers to refer to elementary books on statistics as and when required.*

The most useful device for gathering information is some form of survey by which one means systematic acquisition of facts relating to a very limited field. For instance, nutritionists may wish to know the daily calorific value of diets among a given population. They would thus not be interested in anything that does not relate to this problem. But in order to achieve a meaningful interpretation of their data they may have to include a large number of questions that are not immediately related to eating. So the first effect will lie in designing a questionnaire for this purpose. It may also be necessary to paraphrase the questions according to the special environments in which they are asked, since the vocabulary used to describe symptoms of illness will be affected by culture and social class. For instance a Nigerian farmworker will express himself differently from a British housewife.

The trouble is that in personal interviews the personality of the interviewer may bias the results. One can minimise this by briefing all interviewers very carefully indeed, by precoding interviewing schedules as far as possible, and finally by employing statistical procedures in the evaluation of data in order to compensate for what is called observer errors. In order to get the best out of personal interviews it is necessary to make sure that investigators are very familiar with the questionnaire used so that they can move back and forth through it as information

* An extremely useful reference of this nature is Connolly T.G. and Sluckin W., *An Introduction to Statistics for the Social Sciences*, Third Edition (1971).

is elicited in a relaxed, informal atmosphere which is most successfully achieved by not clinging slavishly to a sequence of questions which are, at best, somewhat arbitrary. In such conditions there is also some room for what is known as "open ended" questions which allow the answer to be phrased in the words of the personal statement of subjects. But as such data are difficult to process and often miss the point of the question, it is advisable to keep them to a minimum. The form of choice is always a precoded question, e.g. "civil status", please ring appropriate answer

married	1	widowed	4
single	2	separated	5
divorced	3	other, please state	

It is essential to allow for the unforeseen with an "other" category if one wishes to avoid suggesting answers. In our example, "other" might conceivably be polygynous marriage, concubine etc. Again it cannot be stressed enough that the quality of the response depends on the clarity of setting out the questionnaire, including detailed instructions of what subjects are required to do and giving an adequate explanation of the purpose of the enquiry to elicit interest and co-operation.

In some circumstances people do not remember past events clearly, or may for psychological reasons be tempted to be inaccurate in their replies. It is possible to incorporate measures for testing internal consistency by asking about the same subject twice. In our nutritional survey we could have thus asked a question on page 1, "How often do you eat meat?"

every day	1
every other day	2
twice a week	3
once a week	4
less frequently	5
never	6

and then some time later ask, "Did you last eat meat on the following days last week?"

Monday	Yes
	No
Tuesday	Yes
	No

etc.

If the answers to the two questions contradict one another it is then possible to allow in evaluation for unreliability of data. From all this it is evident that it is by no means easy to get a questionnaire just right, and it is usually advisable to run a pilot survey of about 1/10 of the final sample size to discover any snags, as for instance that questions on income are not readily answered or that a question is not understood and needs amplifying or that some extra data need to be ascertained. Once a trial run has been made and evaluated, the survey can then take its course and the Pilot data, suitably amended, can be added to the main part since it would not be sensible to include the same respondents twice over.

Having thus finalised the questionnaire which *mutatis mutandis*, implies that the investigators have not only identified the problem area that they wish to study but have also formed a hypothesis * which they will test by means of this survey, the next question to settle is the size and nature of the sample that is to be used. The size does depend very largely on the degree of accuracy that the data is to yield, and statisticians ought to be consulted on this point. In general the bigger the sample the simpler the interpretation of the survey findings are since there is increasing approach to normal distribution curves with large numbers of observations. And also if one deals with really really large populations running to several thousands, then even findings that are not statistically "significant" can be spoken of as meaningful trends. The factors that will control size on the other hand are cost, availability of interviewers or staff and machinery to handle data. Frequently, of course, the size is dictated by the number of affected people as e.g. an enquiry into diabetes cannot exceed the sample size set by the number of diabetics who can be identified in a given area.

So in medical surveys the sampling frame is often very simply one of patients and matched controls, and the skill lies in the "matching" which may need to consider such variables as age, sex, and occupational background, sibship, marital status, parity, area of residence, tribal affiliation, education etc. according to the type of problem to be investigated. In surveys of a more sociological nature sampling techniques must take care to be entirely random, which means that they must not be biased as could happen when restricting interviews by residential areas so that maybe only poor or only well-to-do people are respondents. Other biases that may be introduced arise by having too many women in the sample population which may happen in a survey undertaken during the daytime when men are out at work. This could also lead to a bias in favour of older people. Samples must also be representative by which we mean a population which closely resembles the overall population composition of the area investigated. Thus if a survey is to have national validity it must show the same proportions in its demographic variables as are to be gleaned from national statistics. Sometimes certain areas are atypical, for instance an industrial centre in an otherwise rural population, and then one must either correct for this or state the limited applicability of the findings.

There are various techniques of collecting random samples which are representative of defined populations. One can, for instance, take a map of an urban area, draw a grid of inch by inch squares over it and then decide to include all households in every third square, the result of which is to give a sample roughly 30 % of the town. Alternatively it is possible to interview every third house, but then allowance must be made to compensate for empty plots etc. In Market Research the most

* It is conventional to phrase this as a *Null* hypothesis, i.e. to test how far the data departs from this. For instance, if in the nutrition survey we are interested in the relation between sugar consumption and diabetes, we would test how far the data disprove the notion that these are independent variables.

common technique used is stratified random sampling, which determines the target size say 1000 persons and then splits the population into groups to accord with national or local characteristics, thus for example including.

x	females	25-50	in social classes	1 and 2
Y	"	25-50	" "	3
z	"	25-50	" "	4 and 5
	"	50 +	" "	1 and 2
	"	50 +	" "	3
	"	50 +	" "	4 and 5

and similarly for males, the required proportions having been ascertained from available statistics. Criteria for stratification will, of course, vary with the type of survey and may include political or religious affiliation, length of education, level of income etc The number of variables for which a sample population can be stratified is of course largely dictated by the overall size. Thus in a very small sample it may be possible to represent only one variable with any accuracy, and this can lead to systematic sampling errors.

Problems obviously arise when there are little or no general demographic data available and this is especially likely in rural areas of third world countries. In this case, it may be necessary to start by producing a sample frame i.e. to identify a defined population, before detailed investigations can be meaningfully interpreted against the population background. Such a sample frame has to be built up from general demographic census type data, the accumulation of which is a laborious and painstaking job which needs to be kept up-to-date continuously by adding births and newcomers and subtracting deaths and emigrants. This type of information is the cornerstone of an meaningful medical survey as well as an absolute essential for social planning. Moreover, such detailed records will be of immense value in helping to re-build a community in the aftermath of a disaster.

In industrial countries such statistics are readily available. In England and Wales the annual returns of the Registrar General for instance provide a mine of information on demographic changes as well as on causes of deaths by age, sex and locality. There are, moreover, very detailed assessments made in decennial censuses so that it is often possible to produce revealing medical studies by an analysis of official statistics alone. Although the gathering of statistical information on this scale is clearly impossible in some rural communities, nevertheless attempts in this direction may be of great help in medical or social strategies.

Inasmuch then as modern approaches are scientifically biased, the need to acquire some familiarity with quantitative methods is obvious. It may sometimes be disputed whether we really need measurement in the social sciences but if the investigator is able to keep the problem in perspective, quantitative statements may help in an appreciation of the extent to which statements explain. For instance, when an author reports the incidence of cross-cousin marriages or divorce as being frequent, we give far more meaning to this if we are told that one in ten marriages does, in fact, break up. The word 'frequent' may be right in respect of two very different events (the chances of a cousin being available for marriage must be smallish while every marriage has presumably a 50-50 chance of being a success or a failure) but the implications become very much clearer with the aid of a quantitative statement.

Of course, in the area of socio-medical investigation, investigators will be well used to the numerical approach to problem solving, since in modern medicine, clinical trials demanding most sophisticated use of statistics are important research tools. The main difference in applying surveys in the field lies in the emphasis that must go into questionnaire design.

Department Social Studies
Oxford Polytechnic
Headington
Oxford

Renate Barber