

INTERPRETATION OF DATA

*Hans Raj**

I

IN EVERY advanced country of the world, basic data is collected from time to time both by governmental and non-governmental agencies about problems which strain society and government. For this purpose, surveys are organised and qualified and competent social investigators put in the fields to approach respondents in a bid to collect information. Such surveys can be limited as well as comprehensive. Where field surveys are not possible, questionnaires are circulated and respondents provide data through their responses. A lot of data is collected through old records, printed material including newspapers, journals, books etc. and also by observation methods.

Of course, there are several methods of data collection but, while dealing with problems and presentation of interpretation, our main concern is bound to be survey research which is most commonly used in investigating politico-legal problems. Whatsoever may be the method or source used by the social researchers for collecting data, that in itself does not speak or convey anything. It is no better than a heap of papers. However, it comes out with solid facts, speaks volumes, becomes an eye opener for policy makers and awakens many from their slumber when the interpreter puts life into it through his interpretation. Data collection is the first and foremost step in the process of data interpretation. It is followed by data analysis before the interpreter begins his difficult task.

Data collection, analysis and interpretation have always been a hard exercise, time consuming process and a burden on economy, but at the same time—politically and socially—a necessity. Problems relating to data collection, analysis and interpretation are numerous and increase manifold in a developing country like India where finances are meagre, and the number of qualified and responsible investigators willing to collect data objectively limited, but socio-economic and politico-legal problems which need to be investigated innumerable. These difficulties

* Deputy Director, Indian Council of Social Science Research, New Delhi.

become more knotty for any social researcher who takes upon himself the difficult task of data collection due to vastness of the country, people's reluctance to provide correct information, particularly in the remote rural areas, illiteracy and lack of proper understanding of importance of research in policy making process. A lurking fear that information provided might be misused by governmental agencies to their disadvantage has always proved a great deterrent in undertaking this exercise. Lack of publicity in the area of the importance of the problem to be investigated, differences in rural-urban orientation towards research, personal behaviour of field investigators and their inadequate training before facing respondents etc. are some other problems in this regard.

But the most serious problem in the whole exercise of data collection is that of subjectivity of the investigator which invariably gets in, though it is most desirable to keep it as far away as possible. This subjectivity may arise due to the nature of subject matter, socio-economic and political philosophy of investigator, his emotions, group culture, ethnocentrism, love for prevalent customs and notions, influence of vested interests in the study, pressure groups etc. But whatever may be the reason for subjectivity, it finds expressions through several ways, e.g., posing or wording questions in a way that the respondent responds in the manner the investigator wants him to respond. Similarly, data may be collected from a universe which has no representative character, and thus from the very beginning bias may be introduced in the data. Once general conclusions are drawn on the basis of data, obviously many difficulties crop up at interpretation stage. Problems at data collection stage usually arise due to ambiguity in terms and phrases used and absence of clarity in specification of variables.

Another difficulty with data could be that it may not have been collected with sufficient foresight and planning, and thus, when actual stage of analysis comes up it can be found either wanting in information or too massive to be weeded out before any meaningful analysis becomes possible. Still another problem at data collection stage can be that of ignoring or undermining information not easily available though vital to the problem under research.

Data might have been collected at a time when respondents were not in a mood to provide information; much of personal information might have been sought which the respondent was unwilling to give, or desired information might have been collected in a group situation where the respondents were influenced by each other's responses.

A lot of data on socio-economic and politico-legal problems in India has already been collected in the past and continues to be collected even now, but much of it is lying in stores, unanalysed and uninterpreted. Data is understood to mean all the relevant material, past and present,

servicing as the basis for study and analysis.¹ Data collection and interpretation is both continuing and on-going process. But analysis and interpretation can prove most useful if done without any loss of time.

II

After data collection, the next stage for interpretation of data is that of analysis. It has been said that since interpretation is often inextricably interwoven with analysis, it becomes a special aspect of analysis rather than a distinct operation.² But systematic analysis can be possible only when whole body of collected data is at hand. It is the function of systematic analyst to build an intellectual edifice in which properly sorted and sifted facts and figures are placed in their proper settings and consistent relationships so that general inferences can be drawn from them. Data contains several facts and figures, but what needs to be remembered is that these are not free and equal but have many dependents of varying complexities, sources and structures. The facts are never simple because they involve subjective and objective elements in varying degrees and combinations. Analysis also demands thorough knowledge of one's data because otherwise analysis is likely to be aimless.

But before analysis is done it is very essential that data should be critically examined, keeping in mind both aims and purposes of study. It should be so analysed that it becomes easy to interpret, reproducible, readily disposed to quantitative treatment, significant for some systematic theory and possibly a basis for broader generalisation beyond the immediate content of the facts under study. A blueprint of the study based on assembled facts will make analysis both useful and purposeful. Preparation of such outlines should not be undertaken when only scattered facts are retained in mind but attentive thought and care should be given to this blueprint before analysis. If during the course of analysis certain unplanned findings crop up which do not have any 'relevance to the study in hand, these should not be thrown out because what might have appeared at first as extraneous information might later become the key to the understanding of difficult situation.

In data analysis and subsequent interpretation a great deal of trouble can be avoided by making explicit plans for the tabulation of coded data before constructing the analysis outlines.³ All the variables in terms of

1. P.V. Young, *Scientific Social Survey and Research* 10 (New Delhi, 1975).

2. Claire Seltiz, Lawrence S. Writeman and Stuart W. Cook, *Research Methods in Social Relations* 459 (Holt, Rinehart and Winston, Inc., New York, 1976).

3. Darwin P. Cartwright, "Analysis of Qualitative Material", in Leon Festinger and Daniel Kantz (Eds.), *Research Methods in Behavioural Sciences* 455 (New Delhi, 1970).

which the content is to be coded should be carefully listed. A good outline should contain provision for coding name of the study, the number of each enumeration unit, the name of each coder and any other information relevant to tests of reliability or other statistical treatment. Data likely to be divided should be divided in broad categories which are mutually exclusive and each category should be sufficiently exhaustive. When the analysis outline has been completed with all categories defined, a manual of instructions for coders should be written giving these definitions in clear operational terms.

Sometimes a confusion is created between methods of collecting data as well as data analysis. An analyst should clearly distinguish between the two from the very beginning to avoid subsequent problems in actual analysis and interpretation. The method of data collection determines what information data contains whereas the method of analysis defines this information and helps in its location in the table. Such a distinction is also necessary for understanding the relationship between inferences drawn from different studies.

For constructing an analysis outline some steps usually become unavoidable for suggesting clusters of interrelated decisions which the analyst must make. These are that data must be specified or in other words he must be clear what data are required by his total research design. Then he must make explicit plans of coded data before constructing the analysis outline. Although the variables and categories of the outlines will not usually be different for different methods of tabulation, their arrangement within the outline and the system of notation employed in coding may well be quite different. It will be useful to list variables in terms of which the content is to be coded. In listing the variables to be included in the outline care should be taken to ensure that all information needed on the punch-cards is placed on some variables. There are, of course, many systems of categorisation for any variable but the one chosen will depend upon the objectives of the study and the type of measurement being undertaken. But the category chosen should be neither too broad nor too narrow.

Coding of data collected is both unavoidable and imperative both for analysis and interpretation. But many difficulties usually arise when actual coding begins. Quite frequently data collected does not supply enough reliable information for dependable coding. "A careful examination of the data as soon as they are collected and, if necessary, a systematic questioning of each interviewer or observer will avert many coding problems."⁴ It is always better and desirable that each observation and interview schedule on which data has been based should be checked

4. Selltitz, *supra* note 2 at 475.

for correctness, comprehensibility, consistency and uniformity before whole set of data is analysed and interpreted.⁵

III

But once the data has been collected and analysed and huge resources put into it, it will not be socially useful unless properly interpreted. The purpose of analysis is to summarise the completed observations in such a manner that these yield answer to research questions. It is the purpose of interpretation to search for broader meaning of these answers by linking them to other available knowledge.⁶ In other words, interpretation is the search for the broader meaning of research findings. This search has two major aspects. First, through it an effort is made to establish continuity in social research by linking the results of one study with those of the other, and second, it leads to establishment of explanatory concepts. Interpretation is thus the essence of whole exercise which is made by the funding agencies of social research. Needless to say that wrong or biased interpretation can create many havocs with the problems investigated and result in wastage of material and mental resources which no nation can afford. It is objective interpretation based on sound data objectively collected that can be socially useful and serves the purpose for which whole exercise is made. However, no universal rules can be formulated for interpretation. Every interpretation has statistical description and inferences. But before undertaking the task of interpretation it is to be ensured that the whole lot of data has been scientifically arranged. Since reliable inferences can only be drawn when the statistics are strictly comparable, all heterogeneous statistics should be brought to common denominators. Inappropriate and incomplete statistics are dangerous for interpretation and should be avoided.

Since figures can be played with by the interpreter to suit his convenience, proper care should be taken before accepting interpretation of the investigator, particularly when it is based on statistics. As the task of interpreter is highly technical it should be entrusted to highly skilled investigators, because it is likely that in the hands of inexperienced persons even the most reliable data might be wrongly interpreted thereby leading to wrong and confusing results.

Another precaution which needs to be taken in interpretation of data is that this task should usually be given to those who are associated with the project from its very inception. It is hoped that such investigators would know the stages through which the project has passed, the difficulties involved in data collection and the extent to which bias was

5. *Id.* at 475-76.

6. *Id.* at 455.

introduced at data collection stage and thus the extent to which data can be dependable and interpretation reliable.

Data in social sciences cannot be interpreted with the same perfection as that of natural sciences and as such each interpreter will have to proceed with the basic assumption that it has many errors which have crept into it at collection and analysis stage, and he should give proper weightage and account to this aspect while interpreting data. Similarly an interpreter ought not to forget that the averages do play their mischief and thus stand in the way of proper interpretation. So also on the basis of the limited data collected, generalisation should be avoided to the extent possible. In the words of Darwin P. Cartwright, "Generalisation to a limited set of data to a more inclusive universe cannot be made legitimately unless certain procedures followed."⁷ In this regard it needs to be remembered that wherever specified conditions obtain, specified consequences will follow. Generalisation can be made to an extent only when sample is representative. In principle a satisfactory system of sampling materials will consist of four elements, namely, specification of the universe to which generalisations are to be made, a guarantee that every unit of the universe has a known probability of inclusion in the sample, a procedure of sampling which is independent of correlations among the units of the universe, and a large sample to provide a sufficiently small random error of sampling.

Processing of data in itself passes through several stages. The first step in the process is that of editing the data in which the relevant is to be sifted from the irrelevant in such a way that its basic character is not sacrificed, no prejudices introduced and no personal considerations allowed to creep in. For the purpose of objectivity it is always better that those who are personally affected by the findings should not be associated with this task. Then it should be ensured that whole data is uniform and can be subjected to comparative study and for this purpose some uniform standards need to be maintained. It ought to be properly tabulated according to characteristics and attributes of the plan. A good interpreter should always take his target group into consideration. He should come out with his interpretation in a manner that his findings are easily understood and followed by the category of persons for whom these are intended. In other words, he ought to keep in mind standard of understanding of his group. In addition, his interpretation should answer very clearly, precisely and without confusion the questions raised in the study and for which data was collected and analysed. He should provide ample information for the consumer of data to either accept, reject or partially accept original hypotheses. Interpretation should also, to the extent

7. *Supra* note 3 at 449.

possible, foresee the things and discuss future perspectives of the problem under study.

An interpreter should also clearly spell out the scope of the task and the manner in which he has dealt with it, *i.e.*, he should satisfy the audience by clarifying:

- (1) The intention of the study;
- (2) the hypotheses;
- (3) the sources through which data was collected and difficulties experienced in the exercise of data collection, both in regard to man-power and other resources;
- (4) the manner and the sense in which each term used in the data has been understood and interpreted;
- (5) the manner it has been ensured that data has been made uniform and subjectivities introduced at data collection or analysis stage have been weeded out;
- (6) the techniques followed in the conduct of whole study, and, whether at any stage were these at all abandoned, and, if so, under what circumstances;
- (7) the extent to which data collected was defective and undependable and accordingly interpretation to that extent unreliable;
- (8) the tools of data interpretation used by him and the extent to which they were a success; and
- (9) his own limitations and the extent to which he has been unequal to the task handled by him, *i.e.*, his subjectivity in interpretation, his capacity of understanding the whole problem and his falterings in interpreting the terms used *etc.*

An interpreter should not be satisfied simply with the interpretation of the data placed before him but he should make useful suggestions for future research on the subject, if undertaken. But these suggestions again should dispense with subjectivity. In this regard he should act as an honest broker of information. He should be a scholar responsible for the dissemination of knowledge based on available data.

IV

Interpretation of data being almost last stage of exercise of investigation of research problem, it is essential that it is done in a way that it can be easily understood and followed both by laymen and policy makers. Unless that is presented and made available in a proper form all labour will be lost as an exercise in futility. The idea of interpretation is not to test knowledge or wisdom of the interpreter but to ensure that the

society is benefited by the resources put in the exercise. This interpretation can be both descriptive as well as graphic in presentation. It can also be made available to the society in the form of charts, diagrams and tables etc. But the simplest form of representing research findings is the frequency, distribution or tabulation.⁸ Where help of charts is taken for giving interpretation of data it should be seen that these are neither complicated nor overcrowded. One chart should not contain too many facts and each graph on the chart should clearly specify the purpose and be prepared in an appealing way. It should be properly shaded, clear, concise and simple, and scale legend should be very specific. Absolutely necessary coordinate lines should be used. Plotting points on curves used in the maps should be clearly indicated and grid line drawn relatively slightly.

Interpretation can be made presentable with the help of bar charts line charts and pictographs. It can also be segmental presentation. Bar charts can be both vertical as well as horizontal, but when comparative figures are to be shown, group bar charts should preferably be used because it then becomes easy to understand a complex problem with the help of legends, and comparative data can easily be understood and followed by the target group. Since bars and columns play an important role in the preparation of charts which in turn help in interpretation, it is essential that bars should be arranged in some systematic order, be of uniform width, and have suitable titles. Stubs for various categories of a bar chart should be clearly indicated to the left of vertical baseline.

Sometime an interpreter takes the help of pie chart for clarifying his viewpoint. It is a chart which is in a circle, which in turn is divided into sectors, each representing various parts of the whole. But before preparing such a chart it becomes essential to divide data into ratios and percentages.

There are also component element and silhouette charts which are also used by interpreters to make their findings clear. In the case of the former stress is laid on subdivision of geometric form in relation to the whole whereas in the case of the latter changes in a particular direction over a time can be plotted. But only one series of data can be plotted on this type of chart. There are area and volume charts as well, but P.V. Young is of the view that "as a rule diagrams involving area and cubic comparisons should be avoided, because instead of simplifying the comparisons desired, they are likely to confuse it."⁹ Among other charts mention may be made about organisation and flow charts, ranking charts, time series, ratio, frequency and semi-logarithmic charts. With the help of last mentioned chart it becomes easy to show both relative

8. William J. Goode and Pual K. Hatt, *Methods in Social Research* 343 (McGraw-Hill Book Co. Inc., New York, 1952).

9. *Supra* note 1 at 401.

and absolute changes. These combine the advantages of both the natural and percentage scales without the disadvantage of either.

Interpretation of data can easily be made understandable in some cases with the help of pictorial units. But while developing such units, care should be taken to ensure that symbols used should be self-explanatory and represent a definite unit of value, that number of facts represented should be kept at minimum, and that graphs should give overall picture. No unnecessary attempt should in any case be made to pictorially present such facts which cannot be presented with the help of pictorial units.

Maps are an other useful instrument for interpreters of data. "These are often indispensable in locating problems, verifying hypotheses, analysing data and discovering hidden facts and relationship."¹⁰ Quite often an interpreter takes the help of case map with which it becomes possible for him to portray both physical and social environments on the map.

This map facilitates as well as elucidates social phenomenon and provides framework of primary data which is superimposed on it. Then there are five different types of spot maps. In all these, portraying of geographical location or frequency of social phenomenon is done with the help of dots, discs and squares etc. Then there are cross-hatched maps which help in interpretation of rates, ratios and relative frequencies. The principle underlying cross-hatching technique is to arrange the density of lines or stippling in such a way as to give an optional effect from light to dark in intensity of tone and pattern.¹¹

Of course, system of interpreting data with the help of diagrams helps understanding complex problems easily, and those interested in remembering figures can do so without taxing their memory; but the system has its own problems as well. Exact measurements cannot be easily possible, as considerable time and accuracy is needed in preparing diagrams which at times it is not possible to do. Similarly with the help of diagrams no complex and multiple type comparisons are possible.

Quite a lot of data collected can be interpreted and is in fact interpreted with the help of tables, both simple and multiple. The tables are a means of recording in permanent form the analysis that is made through classification and of placing in right position things that are similar and should be compared. The tables are both simple and complex and are one way, two way and multiway. With the help of each table, usually it becomes easy to understand complex data, particularly for a comparative study. But in tables no description can be given, even if that is badly needed. A little carelessness in preparing tables can, of course, lead to wrong conclusions, not good for a healthy social research.

10. *Id.* at 417.

11. *Id.* at 427.

Collection, analysis and interpretation of social data is complex and interwoven process, because terms used are value loaded, data collected biased and concealed, and interpretation usually not free from bias. If the respondent, on the one hand, is not prepared to come out frankly with what he really knows due to various personal and other reasons, the interpreter, on the other hand, is usually not in a position to weed out biased data and base his interpretation on non-value loaded data. In social research, however, these limitations which are inherent in human beings have to be tolerated and accepted and interpretation understood with proper discount for human limitations, failings and falterings.