

APPENDIX

Some Preliminary Notes on the
Inadequacy of Official Statistics for the Social Sciences (Excerpts)

Giovanni B. Sgritta and Adriano Varotti

In the development of sociological knowledge, the researcher runs into considerable difficulties due to the inadequacies of official statistics... Statistical sources are inadequate to solve the daily problems confronting researchers who are unwilling to renounce a method of analysis based on the optic of a specific theoretical perspective (that is to say the method by which reality is interpreted through the filter of specific analytical categories)... Either at the data collection stage or at the point of data interpretation, science acts as a selective system in confronting reality. Facts take up a meaning only when they are selected and organized according to an exact scheme of conceptual reference... This critique of the dogma of objectivity in observing real facts enables us to establish the first fundamental aspects of our thesis.

- (a) It is ingenuous and wrong to think that it is possible to approach reality without any idea or preconceived theories. Even if we free ourselves from ideas and prejudices in a rational way, we may still be relying on unconscious prejudices or on unconscious dogmas.
- (b) The observation of reality requires a theoretical a-priori, even if frequently we are unconscious of this. It is therefore necessary to make clear the theoretical assumptions which are propædæutics to the observation process.

If one accepts these premises, we may now pass to the question which we propose to call the authoritarianism of statistical information or data fetishism... Data, like commodities, take up a value and meaning independent of the social process which produced it, and therefore its relativism becomes hidden, as does its partial and historically conditioned character.

Let us illustrate this point with an example. Consider the following three propositions;

- (a) Two million farms were in Italy in December 1970;
- (b) Mr Rossi is 178cm tall;
- (c) Electra is the best of the Greek tragedies.

Normally the first two propositions would be considered equivalent in terms of the conditions required for their empirical confirmation. Some doubts could rise in respect of the criteria used for defining and measuring of 'farms', or heights. However the confirmation of the two propositions requires agreement only on these aspects of their content. In proceeding to validate the third proposition however, facts of different type intervene. These concern the plurality of the evaluation criteria or the aesthetic theories that have been implicitly or explicitly adopted in the judgement expressed in the proposition... In other words, the validation of (a) and (b) seems to concern only the procedural aspects of the operations adopted for measuring - it has, that is, an almost exclusively technical nature. In the case of the third proposition, however, the criteria, reasons, choices, and theories which precede the judgment, are unequivocally regarded as arbitrary and subjective.

Now we may ask, can the first two propositions really be considered as logically equivalent? Are they really influenced in such a limited way by the considerations we expressed about the third propositions?

Certainly (a) and (b) have in common that they both express a measurement. The first expresses the number of farms in Italy at a certain moment; the second the number of centimetres contained in the line segment delimited by the extremities of the body of the considered person. Furthermore, for both propositions it is possible to define the reference system: a precise territorial unit in the first case, the body of a well identified person in the second case.

Thus certain common aspects are logically equivalent and consequently the original propositions could appear equivalent too. However such an equivalence is unreal because the levels of specification of their conceptual elements are different. Even if no interpretative difficulties arise in defining the measuring unit in case (b) this cannot be said of case (a). For the word 'farm' does not have unequivocal meaning. Its meaning changes according the point of view taken by the data collector concerning the different agricultural or sociological conceptions.

Hence the individualisation of an object requires a more or less adequate specification of all characteristics defining its content, or a selection among the possible definitions. Such an operation concerns the identification of a criterion, a theory and an ideology. The theoretical assumptions determine the characteristics which are taken to define the object. At the same time they locate the statistical conception and meaning of the object within a wider conceptual scheme... We now examine the consequences for the scientific community of the absence of any theoretical and/or ideological scheme directing the process of collection, elaboration and presentation of data. The first and immediate consequence is that theories and research criteria which are different from those adopted by official organs of information will not find any support in the available data... For these theories would then be obliged to compare themselves with information which is for the most part inadequate and confused. Thus the crisis which characterizes the decision-making capacity of the administrative organs of government and the crisis which characterizes contemporary social thought find, at this point, a common constitutive element.

A second consequence is that information supplied by the official sources can only confirm the basic theoretical and/or ideological assumptions, whether they are expressed or not...

Returning to the main point of this paper, the relationship between the research process and statistical data, we must state our conviction that this is not a relationship between two elements which are logically autonomous and heterogeneous (cognitive view-points versus existential data), but is rather a confrontation between two conceptual totalities each having theoretical form. The first defines the object of knowledge and the assumptions of the researcher's analytical perspective, the second, the theoretical assumptions on which the relevant statistics are based. The second of these is the more problematic, and we shall give it most of our attention.

Within this second aggregate every piece of statistical information is imprinted with a specific theory in structure and genesis (implicitly or explicitly), that is to say it is imprinted on a system of observations, more or less complex, which have cognitive character aimed at the explanation of certain phenomena of the social world.

In other words, the observations carried out by the competent organs through data collection, as well as data elaboration and presentation, are always the result of a plurality of choices which inevitably are directed by a criterion having a theoretical status.

In such a situation, that is to say whenever there is a divergency between the researcher's theoretical perspective and the perspective of the central official organs producing the statistical information, the researcher's task is to utilize the available data like a statistical investigation tool, and to use them in the reconstruction of the image of the reality at which his research aims.

Once the reality image offered by the datum has been made problematic (this is an exclusively theoretical operation), the research must pass from the datum to the modalities and collection techniques, to the classification criteria, to the definitions of the units of analysis, up to the individualization of the theoretical and/or ideological perspectives which are implicit in the logic that supports and unifies all these operations.

The success of this operation of critical correction implies that a plurality of other information must be available among which direct statistical investigations are very important, even if they are carried out on small samples.

More generally, we believe that the formulation of these problems must be one of the principle aims if not the only one at which to direct the social-statistical operator's work and training.