ON THE NEED FOR MORE ANALYSIS AND THE INSTRUMENTS FOR ITS ADVANCEMENT

by Amilni Etzioni

Department of Sociology, Columbia University and the Bureau of Social Science Research, Washington, D.C.

The reasons for the lag in the analysis and verification of information on the social processes in the United States are discussed. Suggestions are put forth to eliminate it.

We assume here that the knowledge available about social processes in the United States—either from social science or from various "lay" sources, in particular the government—is overly discrete and insufficiently analyzed. Later we suggest an empirical procedure to verify this statement. Here we would like to add that analysis is lagging to a very considerable extent. That many more facts are collected than analyzed is not the indication of an analysis gap; some such "over" collection is a necessary redundancy and an inevitable "waste." An attempt to eliminate this would entail the regimentation of science to the degree that it would be totally destroyed. Reference here is to a much greater lag, to a philosophy of science which stresses facts and under-evaluates their analysis, to the abundance of facilities for collecting additional facts and the paucity of facilities for the processing of those we already have, and above all to the procedures of semi-analysis detached from a finalized product and often hindering it. The total effect is for the societal actors to have the information they need but not to know what they need to know, or for their vision to be clear at particular points but blurred whenever they attempt to take in more than a small segment of the societal field at a time.

We distinguish between two positions which are essential and legitimate elements of a philosophy of science; each, though, when pushed to its extreme conclusion, undermines the capacity to know. The empirical position is that, in principle, factual statements about the world (including societies) ought to be maintained only if validated by systematically conducted tests, especially those quantitative in nature. The empiricist position is that theoretical efforts tend to be useless, that which cannot be measured is meaningless, that facts count while interpretations are speculations which tend to weaken the power of the findings. The reflective position is that considerations of the relationships among factors and of the meaning of findings advance one's knowledge. The reflective position is that one can advance our understanding of the world best by reflecting about it on the basis of existing data and qualitative insights into oneself or one's environment. If the empiricist position is anti-analytic, the reflective one is against empirical research. On the other hand, the empirical and reflective positions supplement each other to provide for external validity and internal consistency, the two bases of a scientific perspective of the world.

One could test our proposition that the United States societal knowledge system is empirically oriented in several ways: submit to researchers a questionnaire in which alternative statements describe the four positions defined above and ask them which one most aptly characterizes their position. Or, they may be asked, if given $100,000 to study poverty in city X, which research method they would choose, and how they would divide the funds among the following items: review of literature; preparation of a questionnaire (if any); coding; collection of additional qualitative data; processing of data; semi-analysis; final analysis; writing of final report. And, finally, how much time they would allow for each item. The pattern obtained may be compared to that of researchers in a less empiricis-
tic culture or subculture. We expect that empiricistic researchers will shorten the first and last phases of the process and qualitative methods.

A comparison of their norms (expressed in the answers to the suggested questions) to the actual distribution of budgets and time by the same researchers could further test our proposition. We expect it to show that actual projects are even more empiricistic than the norms of their directors. The reasons for this tendency, if found, may be explained—and the explanatory statements themselves verified—if we examine the granting arrangements.

The existing granting arrangements, we suggest, tend to penalize the analyst and favor the collector of data and the semi-processor. This could be demonstrated by comparing these groupings in terms of their incomes (including summer salaries), the funds allotted for overseas travel and attendance at professional meetings, and the amount of money granted for secretarial help and research assistance. In short, we suggest, empiricists are significantly more rewarded than other researchers, and are given greater access to the facilities they require and greater freedom from other work which may hamper their research.

Moreover, the granting arrangements of the federal agencies and the major foundations tend to generate pressure on the empiricistically-minded to conduct more empiricistic work than even they favor. This could be tested by asking a sample of project directors: how many research projects have you conducted over the last ten years? And of these, in how many did you have what you considered ample time for collection of data? For preliminary analysis? For final analysis? For writing of reports? Which of these aspects was the most shortchanged? The same questions may be repeated regarding a shortage of funds for the satisfactory completion of the various stages.

The research reports of these various projects could be scrutinized by a panel of, let us say, three experienced social scientists, acting as independent judges, to categorize the extent to which the data have been fully analyzed. The underlying reasons for insufficient analysis could be ascertained by interviews with heads of research centers and informal interviews with some project directors and research assistants. Those seem to include the fact that in the United States—unlike most West European countries—full-time academic research is not a fully established career. Researchers cannot readily obtain tenure, obtain the title of “professor,” and be voting members of a university faculty. Moreover, the continued payment of their salaries requires, in the majority of the cases and institutions, gaining a new grant as soon as the old one runs out, or in the very near future. As the lead time of a new grant is about a year, the last period of an old grant—when analysis should take place—is one preoccupied in part by proposal-writing, fund-raising, and psychological realignment to the next project.

Probably much more significant is the fact that in most projects the first and middle phases run for a longer time and cost more than initially was planned for or approved by the granting agency. The period in which time and funds run out is, again, that in which analysis of the data should occur.

Personally, I have little doubt that there are in our environment cultural, organizational, and economic pressures toward the empiricistic approach. I do not believe that such an elaborate testing of this proposition is necessary. A few interviews with the heads of research centers or an examination of their products should suffice. Our main concern is with finding the sources of empiricist tendencies (for which the studies suggested above would be helpful; for instance, they suggest that the attitudes of researchers as well as economic considerations are involved) with an eye to suggesting remedial actions.

Assuming that it is established that analysis is lagging severely behind the collection of information, the following measures may be taken to correct the imbalance. They differ in costs, expected productivity, and the extent to which they themselves require research.

A study of proposals should be made to see if (a) analytic proposals are more likely to be rejected than data-collection ones; (b) budget reductions usually entail a propor-
tionally larger decrease in the funds allocated to analysis, and (e) project directors expect it to be more difficult to gain approval for analytic proposals and items and, hence, cut them in anticipation. If the answer to these questions is in the affirmative, special funds for analysis should be set aside, granting officers should be instructed not to reduce budget items allotted for analysis, and steps should be taken so that funds budgeted for analysis will not be used otherwise. Time tables should allow for a proper period for analysis and writing of the final report. Research centers should be given grants to finance in-between periods—that is, between projects—to allow for proposal writing for new projects without depriving the old ones. A center which supported my work in an earlier period and with which I am familiar invites researchers from other centers for a "writing" year, getting many more books per ten-thousand dollars than other centers. This procedure has some merits in that it provides for analysis where it is lacking, but it deprives the centers where the work was originally conducted of part of the credit. Also, it is limited in scope; most researchers cannot get such a "writing year;" those who do get it cannot count on it; and the facilities given are mainly restricted to a year's salary, which in effect precludes many kinds of analysis.

Granting agencies' staffs may be asked to answer the same questionnaire offered to researchers; if they tend to be empiricist, steps may be taken to alter such an orientation or to circumvent it. While in the natural sciences teamwork is more widely established, in many social sciences the norm of earlier humanistic research still has a strong hold: one scholar is responsible for the product. While he may be assisted and consulted, the title role is infrequently shared. Research may be conducted to see if those best at organizing the collection of data are also the best analysts, and especially if those best at analyzing data are much less effective at arranging for its collection. (Collection itself is already frequently delegated.) Furthermore, in those cases in which projects had two directors who differed in terms of their analytic tendencies, was the final product significantly improved in terms of the balance between these two elements?

This may suggest a sequential team approach, with two project directors for large projects—both working out the design, one in charge of the collection of data and of the preliminary processing, the other in charge of the final processing and the actual writing of the report, each acting continually as a consultant for the other. This arrangement, it should be stressed, assumes not simply division of labor, but division of labor between persons who differ in their skills along the lines indicated. If such arrangements are demonstrated to be productive, this may help make them more acceptable. Still, many studies will be conducted by one director, often empiricistically oriented. Here it would be best if after a given period following completion of a project, the data were released (unless the project director specifically requests otherwise) so that other researchers—specialists in analysis—could work with them.

A special sub-study should be devoted to semi-processing to check that our contention that most studies stop here because they run out of funds, time, and motivation is valid, and to determine what accounts for the small subset of studies which get through to final analysis. Also, this would help to test our proposition that most semi-processing is "detached," using prefabricated categories not relevant to the final analysis and which can be used as "stand-ins" for relevant ones only at a great loss of accuracy: that is, that much semi-processing is more damaging than useful.

If our propositions, as presented so far, are verified, it follows that there are now in the United States large amounts of unanalyzed, under-analyzed, and semi-analyzed data where additional analysis could be productively carried out. Moreover, as analysis is less costly than data collection, the additional analysis could proceed at a lower cost per finding. Hence, more funds should be made available for (a) "analysis and writing years" for project directors who "completed" one or more studies. Some documentation of the potential value of data to be further
analyzed is to be expected, either on the basis of partial analysis already carried out or the quality and integrity of the man. (b) More funds should be provided for secondary analysis by others than the original researchers, for those stored data which are promising (to be tested, as above) where the original researchers provide a release or are willing to participate in “secondary” analysis with others whose specialty is analysis. (All funding agencies should demand a written commitment from recipients of funds that X years after a recipient of public funds ceased to work with a body of data—unless a special case is made that he intends to work with it again—it becomes public domain and a copy is filed in a central archive, such as the William-Menden archives.) After all, even the copyright to one’s own work, unsupported by public funds, is exhausted after a given period. (This will also allow studies to be rerun and help to maintain standards which may be slipping, a topic we do not pursue here.) (c) If most studies do not adequately analyze their own data, rare indeed is the study which systematically relates its findings to those of others, although peripheral references are quite common. The codification of the findings of several studies, often essential for the transition from facts to knowledge, is obviously not the job of primary research. It requires skills and facilities very similar to those of secondary analysis, as outlined above. Hence, the need for such codification constitutes one more mission for, and another reason why, support should be granted to reflective units and men specializing in analysis.

In the field of codification we run into the “mechanical” inventory which provides detached semi-processing similar to the prefabricated categorization of some survey studies. While such inventories are of use as educational tools and limited reference works, when it comes to analysis, to teasing out conclusions, they do not answer the codification need.

(d) Special training for secondary analysis and, more generally, for analysis is needed. We suggest, as a testable proposition, that much more time is devoted in methods courses to collection of data than to analysis and much of the time that is devoted to analysis is used to reinforce the questionable procedures of detached semi-processing. Hence, there is a need to support the teaching of final analysis by the codification of teaching material, seminars for faculties, and other means.

Second, secondary analysis and codification have some special methodological problems which arise out of their need to use data which are already collected, and often have been collected for purposes other than those of the secondary analysis. Matters such as “correction” of the data by return to primary study of a subsample, by the use of “global” qualitative data, the limits of tolerability of error, and the issues raised by finding in secondary analysis that the primary one was invalid, all deserve more systematic and collective professional attention.

(Manuscript received January 2, 1965)

At this point an enigma presents itself which in all ages has agitated inquiring minds. How can it be that mathematics, being after all a product of human thought which is independent of experience, is so admirably appropriate to the objects of reality? Is human reason, then, without experience, merely by taking thought, able to fathom the properties of real things?

Albert Einstein, “Geometry and Experience”