

It is quite possible that training in simple measurement theory may some day be as important in legal training as is elementary exposure to the theory of cost accounting. I am not a lawyer, but I find that the sharply disciplined minds of some of my lawyer friends grasp the essential logic of such problems more quickly and easier than do a few of my less rigorous colleagues with training in psychology or the social sciences.

#### COMMENT BY PAUL F. LAZARSELD†

The social scientist needs very badly at this moment careful analyses of major research endeavors. The two authors have rendered a real service, not only to the lay public and advanced student, but to the research fraternity itself, which needs continuous self-awareness of its procedures. While one cannot but agree with most everything the authors say, there are two points which deserve special attention, and a few words of elaboration.

A survey like Stouffer's should, in part, be looked upon as a piece of contemporary history. Undoubtedly, there will one day be a great deal of writing on the atmosphere of freedom in the mid-twentieth century, just as there has been much on the medieval mind or on the spirit of the Renaissance. The historian of the future, however, will have survey data at his disposal. Although such data, being relatively precise, invites more immediate generalization and interpretation, it is not different from other historical data. If Lord Acton asserts that power corrupts, the only way to prove it is to look at many instances where people came to power and to see what it did to them, under what circumstances it corrupted them and so forth. We would most likely find that power doesn't always corrupt, and that very often when it appears to do so the relationship is actually the reverse. Corrupt people are sometimes more likely to get to power. In a way, we have the same situation when Stouffer finds a correlation between perception of the Communist danger and permissiveness.<sup>1</sup> Under other circumstances and in other historical situations such a correlation might not exist. For instance, the extent to which people are aware of the danger of atomic warfare probably has very low correlation with the extent to which they engage in war-preventing activities; or if an economic depression or some other major issue were to capture national attention, even in regard to Communism, the correlation which Stouffer found might be much lower. Within his own data one could draw interesting inferences as to the variation of the findings. Groups for which the correlation is relatively low either might not believe

† Chairman, Department of Sociology, Columbia University.

<sup>1</sup> Incidentally, I completely agree with the use of the term "permissiveness" instead of "tolerance." As a matter of fact, that term is being used by a Columbia University group in a survey on academic freedom.

that "intolerance" is the best way to fight Communism, or they might have less strong projective tendencies, to mean that they are not as quickly inclined to consider dangerous things they dislike.

A word might be added on additional techniques by which it would be possible to separate the extent to which danger leads to intolerance and the extent to which intolerance is conducive to exaggerating a danger. From other studies we are assured that if we had repeated interviews with the same people it would turn out that the permissive people, even if they considered Communism a great danger, would lose their fear after a while. Conversely, people who don't think that Communism is dangerous, even if they are non-permissive, would on future observation turn out to have quieted down in their restrictive attitudes.

A similar relationship exists between education and permissiveness, and it is undoubtedly important to have psychological theories to explain it. But, it is also interesting to have in this finding a further characterization of the low-education groups in America. We know, for instance, that they are generally less prone to criticize whatever is offered to them in the media of mass communication. They are against federal aid to education. As a whole they have a level of aspiration lower than their actual circumstances necessitate.

These are undoubtedly important social and historical facts, and probably different from what we would find in Europe. If we now find that low-educated people are most restrictive, this adds to our knowledge of the "portrait of the underdog" and permits us in turn to interpret other findings. The tendency to restrict other people could well be the reverse side of a coin on the other side of which we might find a tendency to submit to authoritarianism. For times of totalitarian danger, this could be a significant finding.

Finally, this view of survey data as partly an element of contemporary history also affects the problem of the relation between attitude and action. Again the authors' discussion of this point is convincing. But another element can be added: suppose that McCarthy, after a Republican electoral victory in 1956, would again come into the limelight. Would he find a following? We cannot predict, but whatever the case might be, we will understand the final outcome better if we know how various sectors of the population felt about these issues two years earlier.

Let us turn now to a second point which is of a more methodological nature. The authors point out, correctly, that almost any question in Stouffer's study would permit of a variety of interpretations. Good lawyers would be the first to make this point, but I wonder whether they are aware that this is not just a shortcoming of Stouffer's data. Rather it is inherent in the logical nature of the concept of attitude; and, as a matter of fact, of pretty much every concept used in the social sciences. To tag the idea we could use the term "diagnostic process": whether a man is liberal, whether he has status in the community, whether an armed unit has morale, or whether an educational system is a success or not—all these questions imply that we can classify people or groups in a way indi-

cated by a certain classification. But obviously, one cannot see morale or status the same way as one can discover the size or color of an object. This, incidentally, has nothing to do with the fact that attitudes are something mental or psychological. Whether two men are friends or hostile to each other is a mere matter of external observation, but still it cannot be perceived directly. There are always indicators needed to make an inference as to the underlying classification, just as we need a variety of tests to find out whether a person has a certain kind of illness.

Now it is of the essence of diagnostic procedures that the indicators have only a probability relationship with this intended classification. If a man is a miser, we would find that out by noticing that he doesn't give money to a beggar, that he doesn't loan books and other things. But, there will always be exceptional situations where he will do one of these things and therefore our observations only permit a probability inference. As a matter of fact, we will have to collect a great number of observations to insure that the probability that the man is a miser is high enough to support conclusions or subsequent actions. This has led to the fundamental idea that all such classifications are to be based on batteries of indicators. A typical example is an intelligence test where the classification consists of very many questions because a smaller number would mislead us.

Indicators vary according to the degree to which they indicate what we really want to know. If we want to study a person's anxiety, we might give him first a variety of pictures which are somewhat ambiguous in their content and ask him to interpret what he sees. If his interpretations refer to many dangerous situations, we would assume that the person is quite anxiety-ridden. The inference from the interpretation of such a so-called Thematic Apperception Test to anxiety is certainly more complex and less certain than the inference from a test of mathematical performance to the underlying classification of mathematical ability. It is important to understand this distinction between what is sometimes called *manifest* behavior and *latent* classification. If one wanted to put it facetiously, one would have to say that the answer to the question, "Are you permissive toward radicals?" has only a probability relationship to the intended classification of a man as permissive toward radicals. The degree of certainty of the probability inference from the manifest data to the latent attributes depends upon many factors, one of which is the degree to which a question permits of a variety of interpretations. Although it is a priori impossible to formulate a question in a way that permits of only one interpretation, one should keep in mind that correct interpretation is only one of the many factors which make the diagnostic process uncertain. There are others, such as the respondent's experience immediately prior to the interview. A man who belongs, by and large, to the middle group of permissiveness might give very restrictive answers one morning when his children have irked him a great deal, and much more permissive ones another time when he has just finished a very good meal. To re-

peat: all indicators have only a certain probability with which they are related to an intended classification; the latter can therefore never be reached with certainty. As a result, all classifications in social research have to be more or less impure.

This is true not only for attitudes, but for classifications seemingly much more objective, like the social status of a person or whether a college is church-affiliated or not. As a matter of fact, in terms of social research it would not be possible to classify people with certainty as to whether they were lawyers or not, since graduation from an unaccredited law school, never having had a client, or other such items would be indicators which in individual cases might lead to contradictory results. The reason why this is often not obvious is due to a confusion with normative classifications. If we define a lawyer as one who, by the law of his state, is defined as a lawyer, then we have no trouble; but we have dodged the issue.

In the realm of attitude studies all this has a significant implication with which every research effort has to reckon. It could be described briefly as statistical stability of answers connected with considerable individual turnover. It will be found that most attitude questions which are asked a second time, after a period when people should have forgotten their first answer, show a very great stability. About the same percentage will have said yes to the same question. But sameness here means in the aggregate. Individually it will turn out that a great number of people have changed from yes to no, but they are offset by those who have changed from no to yes. Theoretically speaking, another situation would also prevail: if we were to ask a certain attitude question of one person many hundred times, assuming that we could make him forget his previous answers, we would not expect him always to answer in the same way. He would answer yes only a certain proportion of times, which would be indicative of what one might call his underlying tendency or attitude. The occasional deviations to a no answer would again be due either to variations in interpretations from time to time, or to changes in his mood.

The consequences of this finding for the interpretation of survey data is twofold. For one thing, frequencies and scores in scale values really make sense only for groups of people and not for the individual person. This, fortunately, is no tragedy. In social research we are mostly interested in various groups and the differences in their attitudes and behavior. The same statistical fact is much more bothersome in clinical work, where an individual person is the focus of interest and where, therefore, a much greater variety of instruments have to be used to make what one might call valid inferences.

The second consequence is that most correlations found in social research are notoriously low, and are very likely to stay so. It can be shown mathematically that if tests and other measuring devices have considerable statistical stability, but at the same time considerable individual turnover, then the interrelations between two or more of such measurements cannot be high. This, however, is

not a mathematical shortcoming which we can hope to overcome one day. It reflects the realities of the social world where there are so many disturbing factors and variations that high correlations between a few variables cannot be expected.

If one is inclined to feel pessimistic about such an outlook, one can take consolation in a compensatory aspect. In the social sciences we can take great advantage of what is sometimes called "the interchangeability of indicators." It has been shown, for instance, that different social strata have very different attitudes on economic and political matters. But what is meant by social strata and how should they be measured? We could use an instrument which is based on the possessions which people have, or on their income; we could even use education as an indicator of social status. As a rule it would turn out that whichever index we use, the correlation between attitude and status would be about the same. In other words, the findings of empirical social research are to a considerable extent invariant against reasonable changes from one index to another. There is, for instance, no doubt that the whole classification of permissiveness could have been carried out with a set of different questions, yet the relation between permissiveness and education would have remained practically the same.

The preceding remarks should not be interpreted as pessimistic. The two authors, in discussing the Stouffer study, clearly see where further progress has to be looked for. It will consist of the introduction of ever more sophisticated variables to describe the total structural picture of an attitude. The authors point out that intensity of feeling, information, and other factors have already been utilized. We can immediately add some more, such as the biographic experience of the respondent, his general level of criticism (irrespective of what the object is), his willingness to commit himself at all, and so on. On the other hand, whatever new aspect we introduce will again have to be established by batteries of tests and indicators which have only a probability relation to the intended classification, and therefore are subject to all the limitations of the diagnostic process. But rather than trying in a utopian way to circumvent these limitations, it is better to see all their logical implications and use them in the devising of new research instruments, as well as in the interpretation of research findings.