

Sociology

<http://soc.sagepub.com/>

Problems with Surveys: Method or Epistemology?

Catherine Marsh

Sociology 1979 13: 293

DOI: 10.1177/003803857901300210

The online version of this article can be found at:
<http://soc.sagepub.com/content/13/2/293.citation>

Published by:



<http://www.sagepublications.com>

On behalf of:



[British Sociological Association](http://www.bsa.ac.uk)

Additional services and information for *Sociology* can be found at:

Email Alerts: <http://soc.sagepub.com/cgi/alerts>

Subscriptions: <http://soc.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

PROBLEMS WITH SURVEYS: METHOD OR EPISTEMOLOGY?*

CATHERINE MARSH

It may seem untimely to start worrying about the philosophical basis of survey research at a time when the main difficulty facing any of us is most probably getting cash to do the research at all. However, perhaps for that very reason, arguments, which are declaring that survey research is after all perhaps not on a very sound epistemological footing are gaining currency; I was forced to reflect hard on the process of decision-making in large-scale organizations when the SSRC suddenly discovered in 1975, when funds were beginning to dry up, that it had changed its mind on the importance of survey research and decided to close the SSRC Survey Unit. Many of those arguments and arguments since have made vague references to unease about survey research as a method applicable to producing sociological theory, and some of them I think touch chords in all of us when we consider what contributions to sociological theory have actually been made by survey research.

The trouble is that there are two bogeys to be dealt with at once. The first and most serious is the anti-scientism prevalent in most British sociology today which charges all scientific attempts at the construction of social theory 'positivist', and which holds that technical errors are the result of this philosophical mistake. But the second is the existence of a certain amount of positivist thinking among survey researchers which allows the former confusion to persist.

I am not going to defend the scientific method as such here; that is too large a task. But I hope to relieve it, and surveys as part of it, from the accusation of positivism. I shall do this by examining first what constitutes survey research as a particular type of research method and what are the problems peculiar to it.

I then further want to consider whether these problems are intrinsic philosophical problems which place absolute constraints on the method, or whether they are technical problems which are in principle capable of a solution. It is my contention that, behind the war-cry of positivism, attacks that have been parading as fundamental criticisms of the epistemological basis of survey research have very often been either criticisms of a practical technical nature – i.e. criticisms of bad survey research, which all of us would want to agree with I'm sure – or have raised problems to do with the problem of *any* kind of data collection in social sciences, which stem from the problem that the subject matter of our research is conscious, communicates in a language whose meaning is not capable of unique determination, and is capable of changing very rapidly. This is a problem for any social scientist, from the experimenter to the ethnographer, and is not confined to surveys.

My comments will be restricted to surveys which are designed to provide evidence for particular sociological theorizing. Part of the opprobrium attaching to surveys has come from the fact that their form is similar to that utilized by the public opinion pollster and the market researcher, both of whom are more concerned with predictive ability than with explanation and understanding of the phenomena they study. The polling conception of survey research has often tended to rub off on the sociologist conducting a survey. This has tended to produce two types of results. One response is for the sociologist to treat the subjects from whom she collects information as proxy sociologists, to provide the explanations for their own behaviour that she cannot provide. Another response is that the sociologist acts as something not far removed from a lobbyist, aggregating individual

opinions and presenting them, as if their meaning and importance was self-evident. But those developing social theory also may survey individuals by means of interviews, and collect reports of behaviour, beliefs and attitudes (although as I shall show there is no reason why a survey should not systematically directly observe people); for the academic, the responses are *data* whose role is subservient, to act as evidence for the theoretical end point they are pursuing. For such use, it is important to understand the vital differences between self reported behaviour, behaviour, beliefs and attitudes; many criticisms of survey research have been correct to unmask the illicit assumption that verbal behaviour of various kinds gives good access to behaviour outside the interview. It is vital to realize that in a sociological survey, individuals are approached for information because that is the most efficient way to gain it (it is not always even that, unfortunately). They are usually asked to give reports on their behaviour and their beliefs and attitudes are sought.

Practitioners in this area must be clear themselves about the fundamental nature of what they are doing and the limits and possibilities in using this particular method. Survey researchers have probably all got some vague negative justification for survey research in their heads which amounts to a knowledge that the other styles of research that are open to sociologists are in practice inadequate. For most areas of enquiry, the style of research that is based on experimentation is not possible both on practical and ethical grounds. Yet the 'ethnographic style' of research, if I may so generalize about all those various attempts to apply the method of *verstehen* to small-scale situations by intensive immersion in one area, is somehow not rigorous enough to allow its theories to be subjected to any real constraint in the world beyond the researcher, nor is it capable of producing data over a large enough range of situations to allow the scope or generality of its theories to be tested.

I hope to show in this paper that the drive for rigour and objectivity in our research methods does not commit us to a positivist bandwagon, although I concede that most of the textbook discussions of the subject would not allow one to make the distinction.

The survey as a method of testing hypotheses

I want to define a survey as any enquiry which collects pieces of information, by whatever method, over a range of different cases, and arranges the information about those cases as variables; variables therefore must have the property of providing one unique code for every case. The common strategy for the survey researcher who has collected a case by variable matrix of data of this form is to consider the relationship of the variables, either over the whole of the matrix or in subgroups.

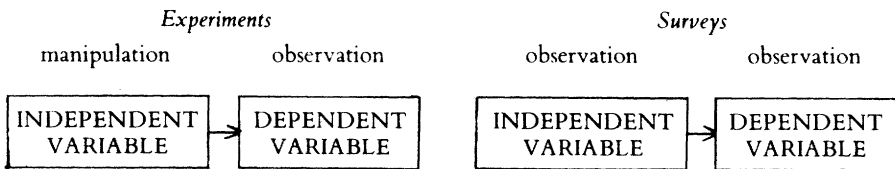


FIGURE 1. *Logically distinct ways of testing causal hypotheses.*

Figure 1 summarizes the two ways known to me of testing hypotheses about social processes. The experimenter 'does something to' her subjects (and usually also 'does not do something to' a set of controls) and looks to see what effect *manipulating* variance in the independent variable has on the dependent variable. Within the limits defined by the laws of probability (if the subjects have been randomly allocated into experimental group and control group) she can be sure that it is what she did to the independent variable that has produced any variance she observes in the dependent variable.

But the survey researcher has only made a series of observations; to be sure, as we shall come on to argue, these cannot be seen just as passive reflections of unproblematic reality, but they must be logically distinguished from the manipulation that the experimenter engages in. The only element of randomness in the survey design comes in random selection of cases; *random sampling is not the same as random allocation into control and experimental groups*. The survey researcher may have a theory which leads her to suspect that X is having a causal effect on Y. If she wants to test this, she has to measure X and Y on a variety of different subjects and infer from the fact that X and Y covary that the original hypothesis was true. But unlike the experimenter, she cannot rule out the possibility *in principle* of there being a third variable prior to X and Y causing the variance in both; the experimenter knows that the relationship is not spurious in this technical sense because she knows exactly what produced variance in X – *she* did. In common with the experimenter, the survey researcher cannot know how X produces an effect on Y; it may do it directly, or it may work through intervening variables.

In other words, in survey research the process of testing causal hypotheses, central to any theory-building endeavour, is a very indirect process of drawing inferences from already existing variance in populations by a rigorous process of comparison. In practice one of the major strategies of the survey researcher is to control for other variables that she thinks might realistically be held to also produce an effect, but she never gets round the purist's objection that since she did not measure everything and control for everything (as the experimenter did by randomization) she has not definitively established a causal relationship.

Furthermore, although having panel data across time certainly helps with the practical resolution of the problem of how to decide which of one's variables are logically prior to which others, it does not solve this logical difficulty that in principle any relationship which one finds may be explained by the operation of another unmeasured factor. And since we are talking about the application of survey research for the elaboration of sociological theory, we can also see that when the subject matter is conscious human beings who are capable of anticipating future occurrences in their actions, knowing that something occurs before something else is no guarantee that it caused it rather than that it was caused by it. (Take for example the relationship between exam performance at O and A level and later jobs taken up among British schoolgirls; there is now quite strong evidence to suppose that the early maturity thesis is quite wrong and that girls, although starting out slightly more capable than boys at getting O level passes, have by A level stage anticipated their later job possibilities and set their horizons lower.)

It is this logical structure which is intrinsic to survey methodology. The data could come from observation, from fixed-choice responses to a postal questionnaire, from content analysis of newspapers or from post-coding taperecorded depth interviews. The important thing is that there is more than one case and that variation between cases is considered systematically. I shall come on later in the paper to argue that the charge of positivism has most bite when applied to some views of structured questionnaires. However, the logical status of a survey is that it is one of two possible ways to test causal hypotheses.

The Essence of Positivism

We are not quite ready to answer the critics who claim that survey research has an inherently positivist bias, however, for we must lay our cards on the table about what constitutes the essence of this philosophy which has almost become a synonym for crassness in common sociological parlance. Kolakowski¹ defines positivism as a philosophy which says nothing about the origin of knowledge but which aims to provide a demarcation

between the knowledge that deserves the name science and that which does not. Although he admits that his intellectual history of positivism has an element of arbitrariness to it inasmuch as he discusses protagonists of this philosophy who with one exception, did not apply this label to themselves, nevertheless he extracts from the wide variation in positivist philosophy four elements which he considers sum up its essence:

- (1) the rule of phenomenalism, which asserts that there is only experience, and which rejects all abstractions be they 'matter' or 'spirit':
- (2) the rule of nominalism, which asserts that words, generalizations, abstractions are linguistic phenomena and do not give us new insight into the world:
- (3) the separation of facts and values:
- (4) the unity of the scientific method.

Obviously all four of these prescriptions have implications if applied to development of sociological theories.

(1) and (2) assert that all knowledge is limited to experience, and that it is impossible to go beyond this to some deeper reality; while historically the development of these postulates about knowledge served a useful purpose in refuting the idealism of the old theological views of knowledge, they deny the possibility of cognitive knowledge. This would lead sociologists to deny the search for underlying personality or social structures which have got dynamics which affect the world as perceived but which themselves are not *directly* perceivable. I stress this: positivism is correct in asserting that evidential criteria have to be sought in our sense perceptions of the world, but incorrect in asserting that theories flow directly from these sense perceptions. Indeed, the opposite is the case: without theory, perceptions are meaningless.

(3) would make sociology a purely technical endeavour collecting enough facts about the world to inform any value position wanted. If we accept this postulate, then the only thing that can go wrong in the process of research is that facts are somehow wrongly reported or perceived. The notion that categories used are inevitably based on a theoretical position and value position is denied by the positivist.

(4) of course asserts the unity of the scientific method on the basis of the first three postulates. Taken out of context it is not objectionable; indeed, I have just argued that there are only two ways to test casual hypothesis in any scientific endeavour. But positivism is wrong in stressing an essential unity of the subject matter of the various sciences, and thus the similarity of causal factors; in the human sciences where the subject matter is social and conscious man, then intentions and motives become an important although not exhaustive component of causes. Moreover, there is no reason to suppose in the social world that any underlying determining social processes *themselves* do not change. This is the major reason why the positivist demand for cumulative empirical knowledge is not easily met.

It is the aim of any science to get at the causal relationships between things; this is precisely what the positivist cannot accept if she buys the first two postulates above, for concepts like necessity and mechanism are abstractions. She is forced to argue that the distinction between causation and correlation is a spurious one, for all that our sense perceptions tell us about is correlation.

Is survey research inherently positivistic?

I stress the word 'inherently' because I would be the first to concede that very large amounts of research in the survey style, as in the laboratory experimental style or the ethnographic style, have accepted the first three postulates of the philosophy that I have argued are unacceptable as the basis of rational knowledge.

NATURE OF PROBLEM

| | PHILOSOPHICAL | PRACTICAL |
|-----------------|-----------------------|-------------------------------------|
| DATA COLLECTION | .g. reactivity | e.g. choice of words for a question |
| ANALYSIS | e.g. casual inference | e.g. level of analysis |

AREA OF
PROBLEM

FIGURE 2. *A typology of 'Problems' in sociological research*

Consider Figure 2 as a summary of the points that I made earlier about needing to distinguish between problems in principle and technical, corrigible problems, and needing to distinguish between data collection and data analysis. The distinction between data collection and analysis is clearly not a temporal one which says that data collection comes before data analysis; this can be illustrated by thinking about the process of coding, which is at once both a method of selecting and collecting information and at the same time is a method of the most primary kind of analysis of that information. I am simply arguing that the problems at the philosophical level of validity, which relate to how to collect valid data and how to draw valid inferences from it, are distinct.

If my earlier argument was correct that the characteristic of survey research was that it was a particular approach to the problem of causality, then a charge of positivism should be found in the problems located in the bottom left-hand box of Figure 2. The argument should be that drawing causal inferences from cross-sectional data, from already existing variation, is unsound because it follows one of the first three epistemological principles outlined.

There are certainly large difficulties associated with drawing causal inferences from snapshots organized in this form, or even from motion pictures: but the procedure is only positivistic if one tries to claim that theory has no role in ordering the variables and assessing the significance of the coefficients. Those who think that there is an automatic way of deciding on the substantive causal significance of a finding through doing a test of significance on a correlation coefficient are certainly wide open to the charge of overt positivism. The question is whether survey research can be used by those who believe that there are processes at work in reality, but not obvious to the observer, to be uncovered by him by means of theory construction and test.

The answer is complex. We cannot order variables in our survey without recourse to a theory about the way the world works, and that theory itself will certainly not have derived from cross-sectional observations. This is not the place to discuss the origins of theories about the world, but in my view they do not spring from passive observations and correlation of attributes of those observations. The results of a survey do not lead to any automatic conclusions about the world which will guide either conservative cabinet ministers or revolutionaries in how to go about achieving their objectives. But they do provide a test (more importantly they often provide the *only* test) of a theoretical hypothesis if the theory is made explicit.

An example to illustrate this

Nichols and Armstrong² in a book called *Workers Divided*, described their depth study of a chemical plant in the West of England and the workers who work in it; unfortunately, although there is much of potential value in this book, it is overshadowed by a complete refusal to discuss methodology. They don't even tell us whether they did their interviews in

the factory or in the pub after work. For two authors who say they reject the survey method (owing to their confusion of the same with fixed format questionnaires) because it is incapable of reflecting in a sensitive fashion the complexities and subtleties of working class thought and ideology, this silence is stunning. Who knows what subtle pressures are at work when the researcher, obviously a committed 'leftie', is buying the drinks?

Although you have to extract the hypotheses from the text with a pickaxe, there are several interesting ideas lurking which could be formalized into hypothetical conditions preventing the development of class solidarity. Taking one of Nichols' chapters from early on in the book, he suggests in effect that one would expect to find a negative relationship between solidarity and

- the management collection of trade union dues
- shift working
- national as opposed to plant wage bargaining
- existence of contract labour on site
- 'massified capital' and complex organizational structures
- management sophistication

These are all suggestions that he culled from his depth study at one particular workplace, **BUT THEY STILL REQUIRE CORROBORATION**. The only way I know of **getting** even partial information to back up these extremely interesting hypotheses is through systematically comparing plants around the country with different situations regarding our hypothesized conditions, to see if there is indeed systematic variation in solidarity. Now just doing a survey of a series of plants and observing variations of this kind certainly would not enable us to draw causal inferences relating the conditions to the development of solidarity, but within the context of a theory which says why these things might be expected to happen, we can use the observations as a limited test. We have to be clear that the test is limited by the fact that it is based on passive observations of the world, rather than manipulation which attempts to change the world in some way.

We may accept a limited version of the Popperian thesis of asymmetry between corroboration and refutation. If we did indeed find a correlation where we expected, even after we had controlled for everything else which we thought was a candidate for **making** the relationship a spurious one, we would still not wish to call this proof of our theory. **But** if we failed to find a relationship at all, we would be tempted to consider this a refutation. I call this a limited version of the Popperian idea because there is always still the theoretical possibility of a third variable acting as a 'suppressor', to use the Columbia terminology; we may fail to find a relationship because sex may be acting to suppress the relationship between shift working and lack of solidarity. It might be that women do less shift work, but are also slower to develop solidaristic class consciousness (there is little evidence that this is the **case**, I hasten to say), and this is hiding an overall negative relationship between shiftwork and class consciousness when you hold sex constant. Nonetheless, I think that we have to say **that** a persistent failure to establish correlation after searching long and hard for possible suppressor variables must be taken as a refutation of the theory which gave rise to the hypothesis.

We are forced to admit that this procedure of inferring causality is fraught with **danger**, and outside the framework of a developed theory is pointless and uninteresting.

Practical problems of data analysis

The example of *Workers Divided* leads us very naturally to consideration of another problem of survey research. Those who have charged it with the accusation of positivism have often found it very difficult to decide whether it committed the atomist fallacy of tearing individuals from their social location and failing to situate them in their social structural setting, or whether the real crime was the aggregative holist fallacy of **thinking**

that there was anything inherently meaningful in characteristics of whole groups or whole societies. Apart from being rather contradictory, there are aspects of both these criticisms that should be located in the bottom right hand box of Figure 2, namely, practical difficulties of data analysis. There is nothing intrinsic to the logic of survey research that dictates the level of analysis, the unit under consideration. The unit has most commonly been the individual, since it is a relatively straightforward procedure to collect information from individuals. Nonetheless, many surveys treat households, firms, or geographical territories or even whole societies as their basic unit of analysis.

Realizing that this is a practical, technical problem, should make practitioners using this method of research reflect very hard about the theory that they are testing, for there are many dangers which surround one when trying to infer features of one level from properties measured at another.³ Hauser⁴ argues that most features of collectivities are, in fact, interpreted as short-hand for properties of individuals, and great care should be taken to measure the original individual properties. He shows, in an entertaining spoof of a respectable-scientific-paper, the dangers associated with inferring a *meaning* to the correlation between the proportion of girls in a school and the educational aspirations of the girls and boys in it. One strand of positivist thinking has historically been associated with holism, and a failure to provide an explanation at the level of mechanism: the mistake of most survey researchers seems to be atomism, however. As Blau⁵ argued in his paper on the methodology of analysing organizations as a whole, 'Quantification, so important for providing evidence in support of generalizations, has often produced an artificial atomism of the organized social structures under investigation'. While I would disagree with Blau's attribution of quantification as the cause (the fact that individuals can speak in answer to questionnaires where organizations cannot seems far more cogent an explanation), I would agree with his diagnosis of the sickness.

The cure is for researchers to think a lot harder about the possible ways in which the distributions in the variables they study might be being generated; Blau was one of the first people to demonstrate that survey research could be used to distinguish between truly individual effects and what he calls 'structural effects', where aggregated individual properties have effects on other individuals. Survey data on individuals can be aggregated to provide contextual measures also.

Let me illustrate this again with the same example of a hypothesized relationship between shift working and the lack of development of class solidarity. If we found a relationship, it could mean:

- (1) that working shifts made individuals less likely to develop solidaristic consciousness, regardless of what other people in the factory did:
- (2) that when many people in the plant were on shift work, the whole plant failed to develop solidaristic consciousness, but that was not more true for shift workers than for others:
- (3) that the existence of shift work in the plant made the shift workers more conscious but this was outweighed by the effect it had on the other staff in whom it produced the effect of complacency.

And of course there could be complex interactions of these effects. Surveys, if properly designed, can investigate the existence of contextual effects of this kind. There is no need *post hoc* to rationalize explanations for how such a negative correlation between shift working and class consciousness came about. If these explanations can be anticipated, with some careful thought and with advice about some of the very complex statistical questions about inference and degrees of freedom associated with different levels of analysis, a study could be designed of several factories and workers within those factories to illuminate some of these interesting questions.⁶

It would appear that the work of the Columbia School sociologists and other European successors, like Boudon,⁷ is little known to English sociologists who are engaged in survey research; and it clearly is unknown to those who argue that there is something inherent in survey research which commits you to a particular level of analysis, even though they cannot make up their mind what that is.

Variables

Finally, we must discuss some of the difficulties which arise from converting individual information to variables and thereafter analysing these either in the whole dataset or in specific subgroups. Blumer's call for what he terms 'generic' variables (I would prefer to call them variables with real definitions) in social sciences is well taken.⁸ But it is important not to blame the process of trying to make one's categories explicit and systematic by 'fitting' reality to variables for the substantive and theoretical paucity of the categories themselves. I have defined a variable as a parameter which has got one unique value for every case and which varies across the population; numbers are usually assigned to the categories in the process of coding, but these categories need reflect no more than a nominal scale of measurement. In other words, all that this criterion for survey research is saying is that when we are talking about characteristics, we must minimally be able to differentiate between the characteristic being present and it being absent. Coding something as a variable 'measured' at a nominal level is doing no more than describing it, making the rules for description in this manner as explicit as possible. Variables are thus simply the result of following through coding rules – they need to be interpreted theoretically before they can be utilized in theory construction. Baldamus⁹ has pointed out that a large proportion of what passes for sociological explanation is merely taking an interesting variable that one would like to 'explain' and correlating it with half a dozen background variables, like age, sex, class, education, religion and so on. This approach results in atheoretical sociology and impossibly boring journal articles.

But the fault is not to be laid at the door of converting complex and rich reality into variables. The fault is the crudity of the way in which things like education are measured, owing to the even greater crudity in the way it is theorized. We do not know whether to measure length of schooling, qualifications, type of institution or attributes of teachers or other pupils, because we have not got a sufficiently specific theory about the importance of various aspects of education. The fact that almost any variable you care to name will produce a zero-order correlation with these background variables reflects the fact that the variables are standing as a simple, miserable proxy for vast and complexly interwoven social institutions.

However, there are difficulties here, because not all the variables are of the same type. Lazarsfeld and Menzel¹⁰ provided an excellent classification for different types of individual properties: these are –

- absolute* properties which are unique to the individual (e.g. age),
- relational* properties which express the relationship of one individual to another (e.g. marital status),
- comparative* properties which derive from a comparison of one individual with another (e.g. sibling order),
- contextual* properties which are formed by associating the individual with the value of his collectivity (e.g. generation).

Causal interpretation of these variable types is very different, since to give a causal explanation is to assert that a change in one variable produces a change in another one, and change in these variable types differs in its implication for other people. If we explain militancy by the absolute variable age, we could just be describing our sample by saying that

the older respondents were less militant. But if we wished to give a causal interpretation of this, we would say that the older a person becomes, the less militant she becomes, and we would also be free to say that if the age structure of the population changed such that mean age increased, then we might expect mean militancy to decrease. But if we think age is standing as proxy for the contextual variable of generation membership individual changes are impossible, but changes in mean value of militancy could be said to have been caused by generational movement. Similarly, if job satisfaction is held to be an absolute property, individual and average changes in it will produce individual and average changes in militancy. But it could plausibly be argued that it is derived from comparing one's own situation with that of others. If one individual's job satisfaction increases, it is bound to mean that someone else's decreases, and thus shifts in average job satisfaction are conceptually impossible and this group causal interpretation cannot be given.

Finally, on the subject of variables, what do we do with those aspects of reality which are important *constants* in human behaviour, which do not vary at all? State power might be considered as something that would fit with the example we have been using of factors affecting the development of class solidarity. Just because something is a constant at the time when we want to measure it, does not mean that it cannot change. Moreover, although state power itself might be a constant itself, it might interact with other variables, especially variables tapping aspects of political consciousness, in producing an effect. Our causal model would be an inaccurate reflection of reality if it did not take this into account.

To sum up, if correlational analysis is used to test theories which link variables in a causal model, then survey research has a contribution to make to the development of scientific theories. We have noted that the idea of cause in our models may not be interpretable in the sense that changing X would necessarily bring about a change in Y, but this is not because the explanation being put forward is not causal. It is because the actual variables being used in the model are not necessarily open to the technical manipulation that would allow the situation to change.

Problems of data collection – the scheduled questionnaire

But we cannot ignore completely the fact that historically the survey method of investigation has been linked with the use of a fixed format questionnaire which is designed so that the transformation of the information on it to computer cards is reasonably fast and straightforward. And it is in the arena of the use of standardized questioning that the charge of positivism has bitten deepest. So I should like to devote the rest of this paper to a consideration of the dangers and difficulties attached to the use of fixed format questions – in other words, of communicating with individuals under fairly controlled conditions. These are questions that affect all sociological research which collects its data this way, not just surveys; you do not escape the difficulties by pretending that you can extract unproblematic information yourself in a pub over a pint of beer.

It is impossible to avoid the problem that asking people questions, as an instrument of measurement, itself 'reacts' upon the person who is being asked the question, and affects the response. Positivism attempts to deny this inherent reactivity; the positivist claims that it is possible to ask unbiased questions, and to get at the truth. This idea of 'absolute truth' lying waiting for a sociologist with keen sense perception and good measuring instruments to tap is absurd. It suggests that perceptions can be atheoretical and value-free, the third postulate. These notions are highly problematic, but they have currency.

In an otherwise sensible book on questionnaire design, Stanley Payne¹¹ defines an unbiased question as one which does not itself affect the answer. What is this supposed to mean? Does it mean that the answer would be an utterance that the same person might have made spontaneously? Clearly not, for any utterance is spoken for a reason, with intention of communicating something: spontaneously there is no reason why people surveyed should

desire to convey this information without a reason. The definition is absurd, for we ask questions precisely in order to elicit utterances slanted in a particular way.

Payne continues: 'One thing has always stumped researchers, and will stump us for a long time to come: having observed different results with different types of questions on the same subject, we still cannot agree on which of the different results comes nearest the truth'. With a definition of truth like this, who wonders at researchers getting stumped? We learn, as we read the (pitifully meagre) literature on question formulation, that you do indeed increase the proportion of people who are prepared to answer negatively to a question by the addition of 'or not' at the end of the question regardless of the subject matter. We learn that the addition of a neutral category in an attitude question which explicitly allows people to remain uncommitted decreases the proportion of those who will endorse the positively phrased items whatever the question. What does this mean? Is one response more true than the other? Certainly not, for different questions were asked: questions are live communications and different questions will convey different intentions of what it is that the researcher wants to the respondent. Our task is to make sure that the intention that is conveyed is the one that we wish to convey, and is not a question about social desirability of something. And certainly every question will not convey precisely the same intention to each respondent, but we shall return to that. We need to know a lot more about what the effect of changing the wording of questions is – we need to know more about the interaction that goes on between interviewer and respondent in the interview situation, so that the interviewer is capable of effectively conveying the researcher's intentions. We must reject Moser and Kalton's¹² prescription of the search for the 'individual true value' (the ITV) which our methods measure with a greater or lesser degree of precision.

What is the implication of this position? Does it rule out the use of the fixed format question? In my opinion, the main conclusion of adopting a position of this kind is that questionnaire design is a very complex task in interpersonal communication, especially if it is designed to stand up to being handled via a third party, namely, an interviewer who did not herself frame the question. It means that before a fixed form for the question can be settled on, piloting various versions of the question and depth interviewing of respondents and interviewers about what they thought the question meant absolutely must occur. Cicourel¹³ does not knock any dents in fixed-choice questions at all by pointing out that the validity of the questions rests on the skills of interpersonal communication of those involved with translating those questions into variables. We must be quite explicit about this and not pretend that the meaning of any of the questions that we ask is self-evident.

But we must also not forget why we bother going to all the trouble of getting a standardized format of communication: it is difficult enough to fully understand the reactivity of this highly controlled situation without multiplying it unduly by changing the question wording also.

Let us return to the question of assuming unity of meaning. Cicourel argues that for a fixed form of question to produce valid answers, the question and the answer would have to be in everyday language not altered by 'particular relevance structures'.

Taylor¹⁴ criticizes the attempt to consider as data 'the subjective reality of individuals' beliefs, attitudes, values, as attested by their responses to certain forms of words'. He believes that questionnaire items are fundamentally incapable of considering 'social reality as characterized by intersubjective and common meanings'.

To the extent that critics in the hermeneutic tradition have made us aware of the centrality of language in many (although by no means all) social interactions, they have performed a very useful function in forcing us to be sensitive to possible ambiguities in the words we choose to frame our questions in and the coding schemes we use for decoding the meanings of the respondents. There is no doubt that there are many sitting targets for this kind of criticism in much social science. The field of opinion and attitude research is notorious for its

blindness to the subtleties of meaning in the questions – it is well known that the general public is strongly in favour of the democratic right to withhold one's labour in an industrial dispute, but draws the line at strikes. There is the ever-present danger of artifacts, and the question creating the response rather than 'eliciting' it.¹⁵

To the extent that criticisms such as these direct the survey researcher towards painstaking piloting of questionnaires, using all the complex skills of a human interviewer to negotiate in a depth interview about the complex meanings involved in respondents' answers, the criticism has been constructive and useful. But, by and large, this has not been the direction of such criticism. The quote from Taylor above illustrates that he believes that fixed format questions can never achieve an understanding of social reality 'as characterized by intersubjective and common meanings', and he is a sufficiently hard-line interpretivist to believe that these common meanings exhaustively constitute the social world.

There are several criticisms one could raise against this point of view. The most obvious is that the social world is clearly *not* simply constituted through language. To be sure, the meaning of the words that we use is inherently problematic. Philosophers who discuss the problems of meaning recognize that the meaning of some words has to be assumed *a priori* as unproblematic so that the meaning of others may be discussed, in order that the 'hermeneutic circle' may be broken. The empirical recommendations for research of authors such as Cicourel who refuse to draw the line at any point over this question of meaning is inevitably sucked into a never-ending circle of negotiation and interpretation. The empirical product of the social science that espouses this view (aptly described by Goldthorpe's characterization as 'DIY linguistics') has not managed to escape from the problems that it has itself identified. And one could argue, against Cicourel, that at one level it is precisely systematic variation in 'particular relevance structures' that we are interested in. A good question will often be one that gets at the appropriate relevance structure, if you like.

Let me illustrate this first with a question about attitudes. If we ask people their opinion on the EEC, it is quite clear that they will not all have the same idea of what the EEC actually is, nor will it be relevant to all of them in the same way. But presumably this is what we would think accounted for differences in their responses to the question; indeed, it is hard to think what else could account for differences, for the EEC itself is not a variable. It makes no sense to talk of response error to a question of this kind. If we ask a question that elicits what has been called a 'social desirability response', we have got to see this as an error in the question, not the respondent.

The danger comes in using words whose ambiguity is unintended and unknown: the meaning of the question to different respondents is varying according to contextual factors that we may be unaware of. But it is important to remember that the most likely result of this ambiguity will be to produce seemingly more random data, obscuring real relationships, rather than leading us into mistaking true relationships which in fact just stem from differences in meaning.

Moser and Kalton, who define response error as deviation from the ITV, admit that there are some difficulties with this conception:

It is true that many questions are not so simple and – for instance – with opinion questions – it would often be difficult to define the ITV. However, this difficulty is beside the point here.

The difficulty is not beside the point. It highlights precisely what is wrong with the positivist conception of truth.

But I might have had a more difficult a time, and Moser and Kalton an easier one, with something 'harder' like the number of rooms in a respondent's house. Here the notion that there is a correct answer that is independent of the question, the interviewer or the respondent, would seem attractive. And yet the post-census survey which checked on the

accuracy of census completion discovered that the definition of a room was not common to all respondents. Some called their landing a room if they did their cooking on it. But my point is that this is a failure in the adequate communication of the intention behind the question, not a 'response error'. The respondent did not err; she merely told the census division something incidentally interesting about the way people define room space. We want the *answer* to vary according to the particular relevance structure, which is here the number of rooms as defined by the census division; but we want the question to invariably communicate the way in which the relevance is to be considered.

Now it may be that there are some intentions that we may have as sociologists which cannot be adequately communicated to respondents. Sennett & Cobb¹⁶ pointed out that sociologists often act as though the syndrome of denial of particularly painful psychological events had not been discovered, for they expect respondents to be able to convey to them the most inner of feeling states. This means that different techniques must be developed, like the semantic differential, which communicate the question at a less conscious level.

We must not confuse an impossible attempt to achieve 'absolute truth' through asking unbiased questions, with the aim of being objective in our quest for truth, through trying to be as rigorous as possible in the way in which we draw conclusions from observations we make about the world, what people say and how they behave: such objectivity stems mainly from making explicit the rules of coding we use.

What is the practical implication of this discussion of reactivity? Most importantly, it means that although studies which shine light on interviewer variance or response instability have got a positive aspect in that they force the research to be aware of the fact that the instrument she is using is a highly reactive one, it is no solution to just use the knowledge to increase one's confidence intervals around one's population estimates as Moser and Kalton recommend. Interviewer variance and test-retest results merely point to the existence of ambiguity through their *net* effects. We must as sociologists be concerned with the whole of the situation, and understand why some interviewers are communicating different intentions to other ones. In order to do this, we will have to investigate thoroughly what intentions they are in fact conveying, and this is something we should be looking at even with stable questions and no interviewer variance. The meaning of a stable response is certainly not self evident, as Cicourel correctly points out.

But, in summary, this is a problem that any researcher who, if forced to collect data in this way, will have to face. Very many experiments have as the measurement of the dependent variable a fixed-choice question to the subjects of the experiment. And certainly most depth field studies advance through the medium of language. These studies do not avoid the problems although perhaps they are much less likely to be able to clearly say the extent of them.

Conclusion

I have been concerned to make a distinction between philosophical problems and technical problems, between problems inherent in analysis and problems in data collection. The purpose of making these is to avoid misidentifying the source of many of the problems that exist in survey research today. Crude data-dredging and false notions of truth and bias have allowed some of the critics of survey research to call the method inherently positivist. We must be clear that there is an alternative and valid way to approach the problem of causal inference and objectivity in social science which survey research can be part of.

We have to clear up these problems in order to tackle the bigger one, which is defending the scientific approach to an understanding of human affairs. We do not need to support a very strong version of the sociology of knowledge to feel sure that funding bodies will be

casting around at the moment for ways to save money, and arguments about the inherent uselessness of survey research will gain an ear. We have to be clear about why these arguments are wrong.

Notes

1. Leszek Kolakowski, *Positivist Philosophy*, London, 1972.
2. Theo Nichols and Peter Armstrong, *Workers Divided*, London, 1976.
3. The classic article here is W. S. Robinson, 'Ecological correlations and the behaviour of individuals' in *American Soc. Review*, 15, June, 1950, 351-57. A more recent discussion of the problem can be found in Hayward T. Alker's 'A typology of ecological fallacies' in Mattei Dogan and Stein Rokkan (eds.), *Quantitative Ecological Analysis in the Social Sciences*, pp. 69-86, Cambridge, Mass.
4. R. M. Hauser, 'Context and consex: a cautionary tale'. *Am. Journal of Soc.*, 75, 1970, pp. 645-64.
5. Peter Blau, 'Formal organizations: dimensions of analysis', *Am. Journal of Soc.*, 63, 1957, pp. 58-69.
6. James A. Davis, Joe L. Spaeth and Carolyn Huson ('A technique for analysing the effects of group composition', *American Soc. Review*, 26, 1961, pp. 215-25) work through a very clear example hypothetically drawn from Durkheim's *Suicide* to illustrate how one could accomplish this. Hauser's influential misgivings (*op. cit.*) about this approach should not stop one using it, but rather spur one on to anticipating the explanations one would give of the contextual effects so that one could design questions to test this.
7. I have already cited an example of Blau's work. See also Raymond Boudon, *The Uses of Structuralism*, London, 1971, and P. Lazarsfeld 'Problems in methodology' in R. Merton, L. Broom and L. S. Cottrell, Jr., *Sociology Today*, New York, 1960.
8. H. Blumer, 'Sociological analysis and the variable', *American Soc. Review*, 21, 1951, pp. 683-90.
9. W. Baldamus, *The Structure of Sociological Inference*, London, 1976, p. 125.
10. P. Lazarsfeld and H. Menzel, 'On the relationship between individual and collective properties' in A. Etzioni (ed.), *Complex Organizations*, N.Y., 1961, pp. 422-40.
11. Stanley Payne, *The Art of Asking Questions*, Princeton U.P., 1951.
12. C. Moser and G. Kalton, *Survey Methods of Social Investigation*, London, 1971.
13. A. V. Cicourel, *Method and Measurement*, New York, 1964.
14. Charles Taylor, 'Interpretation and the sciences of man' in Roger Bechler and Alan R. Drengson (eds.), *The Philosophy of Society*, London, 1978.
15. See Catherine Marsh 'Opinion polls: social science or political manoeuvre?' in Jeff Evans, John Irvine and Ian Miles (eds.), *Demystifying Social Statistics*, London, March 1979.
16. Richard Sennett and Jonathan Cobb, *Hidden Injuries of Class*, New York Vintage Books, 1973.