The Journal of Mind and Behavior
Winter 1981, Volume 2, Number 4
ISSN 0271-0137

Is Field Work Scientific?

Linn Mo

The University of Trondheim

Only a deductive nomological explanation is scientific according to the rules of formal logic. This procedure takes us down a ladder of logic, but includes subjective judgments at the point where we choose indicators. Inductive reasoning is an attempt to go up the ladder of logic and is widely used, particularly in natural sciences and in quantitative procedures such as factor analysis. A logical fallacy in requiring reliability in this type of analysis is pointed out. Both procedures are uncertain ways of gathering knowledge, as the number of theories that could explain the data are potentially endless. Social scientists' disagreement about method reflects lack of a paradigm, but paradigms can also be wrong. The nature of the field sets the important limits as to how scientific we can be.

This article addresses two issues: that of the inductive reasoning used in most field work, and that of precision in data collection and analysis. Field work is of course only a means of gathering data, while science is concerned with answering questions. Not all types of questions are of interest. Questions of fact or questions about some specific situation are not considered relevant unless they have some theoretical import. What we are generally looking for is an *explanation* of the phenomenon to be investigated.

What then, is satisfactory explanation? According to formal logic, a phenomenon should be explained in terms of a cause-and-effect law that predicts it. This is called a deductive nomological explanation. Deductive logic provides us with a ladder from the theory down to the specific instance we are investigating.

The theory is only *proposing* an explanation for the phenomenon. (According to this way of thinking, it is quite irrelevant where the theory comes from. It can be an old one we found in our textbooks or a new one we invented.) Climbing down the ladder of logic, we deduce a hypothesis from the theory and then design research that will *discriminate* among the possibilities. We are testing the theory by constructing a deductive argument with theory as the premise and a hypothesis about a specific phenomenon as the conclusion. In a valid deductive argument, it is impossible for the conclusion to be false if the premises are true, for the

Requests for reprints should be sent to Linn Mo, Ph.D., The Norwegian Institute of Technology, Division of Organization and Work Science, Alfred Getz. Vei 1, N 7034 Trondheim-NTH, Norway.

conclusion is contained implicitly in the premises. Thus if the hypothesis is refuted or not supported, we can infer that something is wrong with our theory. At each point in this procedure, inference is supposed to be logically explicit so that anyone with common sense can agree with the reasoning.

Deductive hypothesis-testing is supposed to be a feedback process allowing progressive adjustment or rectification of the theory. Only the refuted hypotheses help us to do this. Research results that support the hypothesis prove nothing, for another piece of research, testing the same hypothesis in a new specific situation, might subsequently show negative findings. Support, then, is not definitive. In practice, however, refutation is not definitive either. We may have misunderstood our data, or some technical error in carrying out the research may have lead us to a false refutation of the hypothesis. The theory is questioned, but is allowed to stand, pending further developments.

For some reason, social scientists seem to prefer positive findings. Perhaps positive findings are more publishable, or perhaps this is only an instance of human behavior—we like to show that we have guessed right. Thus old theories seldom die, they just fade away or are incorporated in the new.

Another problem is that positive findings can also be explained by another theory. We try to design research to discriminate between the theories we have, but there is nothing to prevent some bright soul from proposing a previously unthought of theory that competes with ours in explaining our findings. This puts science in a rather dismal position. At best, we know nothing except what is *not* the case, what we have found to be wrong guesses. We are doomed to accumulate knowledge by a tedious process of elimination, in a world where the theoretical possibilities are potentially endless, the number of hypotheses that can be drawn from any one theory are potentially endless, and the number of situations where any one hypothesis could be tested and refuted are potentially endless.

Social science has some additional problems in applying this model of what science is all about. Testing hypotheses is only possible where we can manipulate a situation, or find a repeating situation that is convenient for testing precisely our hypothesis without contaminating the test with extra variables. We cannot do this when the situation we are trying to explain is unique (for example, some historical event). When we *could* arrange a test, the manipulation may be unethical, or extremely impractical in terms of the cost, the time it would take, or the number of people who would have to be mobilized. Human conciousness may also in-

terfere with the test. *People* as research objects can react to hypotheses about their behavior, either by making them into self-fulfilling prophecies or by rebelling and behaving differently in order to prove us wrong. Of course there is still a lot of social research we *can* carry out according to the deductive model, but some very interesting questions (from the point of view of the discipline) have to be either ruled out as research questions or investigated a different way.

Inductive Reasoning

A deductive nomological explanation only answers a "why" question, and the answer must be testable against further empirical evidence. We ask other kinds of questions too, especially "how" questions—not all questions can be broken down to "why" questions. Sometimes we want to explain some unique event after is has happened (retrospective causal analysis). Sometimes we just want to watch a whole situation developing, in order to decide what aspect of it is important to social science. In these cases we turn to inductive reasoning, working from the specific phenomenon up the ladder to a theory. We examine the relevant data first, making them our premises. We do not have any prior theory or hypotheses as preconceptions, but come to a theory as a conclusion.

Inductive arguments are not absolutely valid or invalid as deductive arguments are. If the premises are true, it is more or less improbable that the conclusion is false, so we say that inductive arguments are probabilistic, with various degrees of strength. The reason for this is that an inductive argument is always making some claim beyond the premises, either predicting the future on the basis of the past and present, or generalizing from specific situations or instances, or constructing an explanation from incomplete data or clues about an event that cannot be repeated.

Although this method is sometimes called the "essence of science" it has always been controversial, and has some serious logical weaknesses. First of all, induction must add something, not just summarize the data. The resulting theory is supposed to be a new discovery, previously unthought. How do we get there from here? In fact, there are no generally accepted rules for getting *up* the ladder of logic. We have to jump. This jump is variously called a "transcendental leap" or "jumping to conclusions" depending on the quality of the theory produced and on one's point of view.

We do think inductively (or we think we do) but how can we justify it

in terms of formal logic? A number of lame arguments have been tried.[1] Some theorists are content with a feeling of "subjective probability" that convinces them of the rightness of their new theory. Some are convinced by the sheer number of supporting instances. These arguments do not satisfy the requirements of deductive logic. Some get around this by changing the requirements. They say we *can* evaluate evidence for one hypothesis without reference to alternative hypotheses, or that we *can* make rules for decisions based on utility as well as evidence, if the evidence seems a bit flimsy. Others simply reject the rules, denying that they are *any* valid principles for assessing evidence, or dismissing the whole issue as an argument over the meaning of the word "rational."

Some theorists make up sets of rules for the inductive method, but run into more logical or empirical problems. They must usually assume that nature is uniform (*is* it?) or that the purpose of science is to find frequencies of events or probabilities (*is* it?).

In everyday life, we operate on the basis of working hypotheses drawn from previous experience (inductive inference). We are sometimes surprised, but we expect to be surprised some of the time, for although we assume a kind of uniformity of nature, we also allow for the possibility of "freak accidents". This approach serves to justify inductive principles inductively, arguing that they often work, but it does not explain why we make mistakes. Sometimes the working hypotheses drawn from previous experience are simply wrong.

Another approach is to try to find the rules for the method empirically, by examining successful instances of inductive reasoning. Some scientists have made "transcendental leaps" or at least have become famous for new solutions to long standing scientific problems. This "leap" can happen suddenly, perhaps after a great deal of thought about the subject, but without any concious premonition of exactly this particular solution. The new solution has a "wholeness" about it that makes it recognizably better than all previous attempts. Studies of creativity look for causes in the way the ground is prepared, and books about the inductive approach can tell us how to prepare the ground[2], but so far there is no support for the claim of a "logic of discovery". Empirical studies of success do not explain why we make mistakes either. Many experimenters prepare the ground at least as carefully and receive no inspiration, or they find an answer that turns out to be wrong. Real discovery seems to be pure luck (or intuition, creativity, serendipity), and

[1]These are summarized in the introduction to Kyburg and Nagel (1963).

[2]Glaser and Strauss (1967) give field workers helpful hints about how to prepare the ground.

it seems that serendipity shines on the just and the unjust alike. The most that can be said is that sometimes a discovery occurs and sometimes the results are useful. The fact that we do not have any logical rules for finding what we seek does not mean the very idea is absurd.

A second objection is that we cannot really examine the relevant data without prejudice. We always start with some preconceived idea—or do we? Watch the way both people and animals learn. Learning is an exploratory process where repeated experiments lead to a discovery. A small child reaches out to a lovely dancing flame. After a number of painful trials, the child looks but does not touch, or perhaps he or she reaches very carefully, obviously testing a hypothesis in true scientific spirit. We say the child has learned--but what do the previous trials represent? Was the child gathering data without preconceived ideas about fire, or was he or she testing the hypothesis that *it would not hurt?* In such situations, and in much of our own thinking, the boundary between inductive and deductive reasoning is indistinct.

The study of learning does, however, offer us an explanation of mistakes. We do not "see" *all* the relevant data, we "see" only what attracts our attention. This can be another way of saying "luck"—a previously unnoticed piece of data was suddenly noticed. Why was it not noticed before? We could say that it was a technical error—a preconceived theory or hypothesis drew our attention away from it. If we had been applying the inductive method correctly, observing without prejudice, we would have seen *all* the relevant data.

This leads us to another dismal position, for just as there is no upper limit to the number of possible hypotheses to be tested before we accept a theory, there is no upper limit to the number of possible facts to be considered before we propose a theory, and still no upper limit to the number of possible theories that could explain the data. The practice of induction relies on the same unspoken criteria of "relevance" or "importance" that govern the practical application of deductive hypothesis-testing, and *these criteria can be wrong.* In practice, we are satisfied after a number of trials, a number of tests, or a "long enough" period in the field. To go back to the small child discovering that fire is hot, suppose the child persists in testing the hypothesis every time he or she sees fire? At some point we would stop praising the child's scientific spirit and begin to question his or her intelligence, but here we are appealing to subjective judgment to get around a logical problem. Science as a whole is inductive in this wider sense, in that we are satisfied at some point and file a theory as "knowledge."

Finally, one may object to the results of inductive research. The

answer to a "how" question or a "what is going on here" question is usually a description, analytic theory, or heuristic model. This kind of theory is not falsifiable. If we insist that the purpose of science is to discriminate among possibilities, and not just to enlarge the stockpile of uncertain knowledge, then this kind of theory is of no use at all. The opponents and the supporters of inductive methods can only argue past each other on this issue. Our position depends on whether we take it as a philosophical or an empirical question.

Insisting on the deductive hypothesis-testing model as a standard means rejecting most of social science, as well as a lot of what goes on in other sciences, as unscientific. Within social science, the objections to inductive reasoning as a research design for field work must also apply to inductive quantitative methods such as factor analysis and Bayesian inference. If we look at the standards actually used (rather than professed) by other scientific disciplines, we find a good deal of descriptive work and exploratory studies leading to models or probabilistic laws. The models are put to practical use in calculating magnitudes or predicting events. Since they are probabilistic models, the discovery of a negative case is assumed to be possible and does not refute the whole theory.

Most *original* research (in any field) is a kind of playing around. We explore a bit, begin to form opinions, test and reject them, and try a new approach. The report we write at the end of this process does not always reflect the kind of thinking that went into the research. Exploratory work sometimes ends with a testable hypothesis. Scientists in other fields accept this kind of conclusion as Step One, a reasonable point to write up the results. Step Two is testing the hypothesis. Some say you are being a scientist only during Step Two, others say you are being only a technician performing a ritual. Science as a whole depends on both Step One, the apprehension of ideas, and Step Two, the critical voice. Neither is sufficient alone. In social science, it seems that we have different people doing Step One and Step Two, theorists and imaginative explorers on one hand, and, on the other hand, people with technical competence and skills testing theories often not their own. Each group could find very good reasons for calling the other something less than scientists. What we need is more humility about what we are doing, and about what it is even possible to do. We can make educated guesses (that could be wrong) and test them *rigorously* (if that were feasible and useful—they could still be wrong); we risk being mistaken and confused.

The Problem of Precision

Field work could be carried out to test hypotheses, although this is

seldom done. What makes some social scientists queasy about field work is not the inductive approach, but simply the general imprecision of the method. It is not *reliable,* they say.

Reliability cannot be measured in field work. There is little structure for a second researcher to replicate, and besides, each researcher's personality and relationships to the respondents affect the data differently. The researcher is affected by the respondents, and the researcher's presence also affects the situation being studied. This gets particularly messy where the researcher is also *participating* in the situation observed, either in order to understand the data better or out of a feeling of moral responsibility. However, does it help to stay out of the situation and design research that "measures" the important variables? There is a rung missing in the deductive ladder of logic at the point where the researcher must choose indicators. There are no rules for choosing indicators. The question of validity is settled by subjective judgement or by agreement. (This is another instance of the way we jump over philosophical problems we cannot solve).

This kind of research hangs on a chain of interpretations at the bottom of the ladder of logic, and anyone who has been through this process and remembers the conundrums that had to be faced in designing an instrument and coding the data, and the arbitrary choices that had to be made, should have a healthy scepsis about the precision of all those numbers. Of course we expect that the errors are small (although some are unknown or unmeasurable) and the reader of a report is supposed to take account of these. We *feel* more certain because the research, or parts of it, could be done again and the reliability measured (although it could be systematically distorted and the repeat performance distorted the same way). Field workers also have devices to help them feel more certain. Tape recorders, photographs, or video tapes are better than notes, but also affect the participants and thus distort the data. We may have to trade off between certainty and distortion.

Field work lacks precision of the kind that can be measured, but social scientists typically manage to measure only small pieces of social reality, such as indicators, or reliability as an indicator of credibility. Measurement is not the only way, and not always the best way to check precision. A photograph, for example, is a precise record that can only be verified by comparing it with the original scene or by showing it to people who remember the scene as it was. Data gathering through observation and unstructured interviews should be precise in the sense that a photograph is precise. Different researchers may focus on different aspects of the situation, just as photographers may take pictures from various angles,

but the descriptions should be recognizable as records of the same situation, just as we would recognize the same persons or objects in the different photographs. A researcher can be criticized for incomplete or inaccurate data collection, as a photographer can be criticized for poor aim or incorrect focus.

Reliability in the analysis of data poses another kind of problem if we are using the inductive approach. Given any framework, it would theoretically be possible to get an independent coder to read through all the field notes and interviews and make another classification of the details to check the researcher's conclusions. However, this would be much more complicated and time-consuming than re-coding standardized questionnaires with the help of the researcher's code book. Also, what one would find out is less likely to be important, because of the difference here between the deductive and the inductive approach. A code book, and the hypothesis it is related to, are written in advance of looking at much of the data. Then it becomes important to make sure each bit of data has been properly interpreted and tallied. The deductive approach aims to establish a framework and then see if the data fits, while the inductive approach seeks to look at the data first, and then derive a framework. Of course most of the specific findings turn out to fit rather nicely into the framework, because the framework has been derived from a generalization of exactly this set of details!

The situation is similar in inductive quantitative procedures such as factor analysis where the framework is derived from the data. The objective is to find the framework with the closest fit. When an inductive approach is used, the framework becomes the interpretation of the data. The real scientific issue is the original choice of framework. Has the researcher found a true assertion about the real world? Are the most important features of the situation included? These are no longer questions of reliability, but of validity of the framework. Finally, the framework chosen may be valid without it being the only one that is valid. Closeness of fit to the data cannot be measured in a descriptive report as it can in factor analysis, but even in factor analysis the criterion of closeness of fit may be overruled by the criterion that the results must make sense to the researcher so that the derived factors can be named.

If data gathering can be compared to photography, the analysis of data is more like technical drawing, for here the objective is to present the aspect the researcher has decided is most meaningful, a clear "figure" on a "ground" composed of whatever background material is essential to an understanding of the situation. Of course the subject, either described or photographed or drawn, is always unique. All data come from unique

situations, but, in addition, we hope that data gathered and analyzed for scientific purposes can be either representative of a general class of possible data, or useful for showing some general principles in operation, just as a picture of a particular rose can show us what roses are like, or show some structural details common to flowers.

In making the analysis, the social scientist puts a picture together from many interviews and hundreds of pages of field notes similar to the way a biologist would work from many photographs of roses to prepare a single illustrative sketch. The empirical evidence is conveyed together with the meaning the researcher derives from it. The reader verifies the analysis by checking the report against his or her own knowledge of the original situation or similar situations. Of course the reader may have a slightly different analysis, but he or she generally takes into account the researcher's theoretical framework (like the photographer's perspective) and the circumstances of the research, to explain away some of the differences, if they are not too great. Glaser and Strauss (1967) call this a "discounting process"; (e.g., "When I was there it was raining, but this picture was taken on a sunny day.")

The result is not a risky test of a hypothesis, but a risky presentation of one. It could be falsified by others, including laymen who know the situation. All of our theory is interpretation, and however carefully the "facts" (data) may be gathered, the interpretation could be wrong. We are again back to the problem of knowing anything with certainty.

What Can We Do?

Social scientists disagree over what is acceptable method. Some would include intuition and introspection, others draw the line barely to include or exclude field work, some would allow only quantitative methods, some only deductive approaches.[3] The lack of agreement about methods reflects the state of the social sciences, at the "pre-paradigm stage of development".[4] However, paradigms are only theories that have become accepted as "knowledge", or social facts. The way ideas become social facts is not always a rational process in everyday life, and according to Kuhn (1962), it is not such a rational process in science either.

---

[3]Ford (1975) argues for a deductive approach and also discusses the current disagreement in sociology.

[4]Agreement within any scientific discipline is only relative, and changes in degree over time. The methods abandoned after a change of paradigm are not necessarily unscientific. They may be, in terms of the new paradigm, but the new paradigm may also be wrong (Kuhn, 1962).

The lack of agreement about method is also a lack of agreement about the appropriate questions and about the purpose of social science. These issues are all intertwined, as the issues of logic and precision are somewhat interwined in this article. A paradigm would represent agreement on a basic set of theories. If we had this, we would have a purpose—explore the social world in terms of the paradigm. Our purpose would direct us to the questions that were interesting, and the questions in turn would suggest what methods were appropriate.

We do not have a paradigm, but perhaps that is just as well. Paradigms can also be wrong. In the absence of a paradigm, the nature of the field sets the most important limits to what is possible. The things we study (our variables) are not things we can apprehend directly through our five senses. Consider, for example, deviance, social class, power, organization, values, etc. We think we know what we mean by these terms, but the plethora of definitions indicates that perhaps we are not so sure. Social science is the study of meanings. The data are most often people's verbal interpretations of what they are doing, and analysis is building *our* interpretation of these. Our interpretations are neither objective nor subjective, but intersubjective, a more or less shared reality within the field. Of course "less shared" means conflict and controversy, and the danger that "objectivity" (intersubjective agreement) may depend on how many agree and what power or positions these people have. Then of course the whole shaky structure rests on shifting sand as the people we study gradually (or quite suddenly) change their minds about what they are doing.

In this state of flux it is not surprising that social scientists take "positions" on various issues, including research methods. There is room for honest disagreement about the importance of the strengths versus the weaknesses of particular methods. What is remarkable is how often those holding various positions are arguing past each other. Perhaps some positions are unnecessary. Rules, numbers, or ideologically based choices of method regardless of the research question can be the security blankets we cling to in order to forget the real risks we are taking in an uncertain field.

Even without a paradigm we could let the choice of methods be guided more by our purposes and questions. Is our purpose the same as in natural science, to explain and predict? Or is it, *in addition,* to do some things we cannot do with natural science, to understand social life and perhaps affect it? (We cannot "understand" natural phenomena in the empathetic sense, and we cannot affect natural laws, we can only take them into account in what we do.) Our knowledge of human behavior

has little use as predictive power. One observes oneself, as in a mirror, and adjusts simultaneously so that one's image meets one's expectations. Perhaps we can wipe the egg off our faces? As social scientists we can at least try to make good mirrors.

Shall we then limit ourselves to "why" questions, or also admit the "how" questions? Both types of questions are permissable and *necessary* if we are to understand social life. We should ask, rather, whether the question suitable to the discipline is worth investigating in terms of whatever purposes we have, and amenable to some kind of empirical study.

All the extreme positions we read about are illuminating, but we can steer between many of the controversies over method. Are we positivists or idealists? The choice does not have to be made once and for all. The researcher is never either wholly apart from, nor wholly a part of the system being studied. This is a matter of degree, and varies with the project and research question. Shall we take an inductive or deductive approach? This can depend on whether or not we have a hypothesis and whether or not it is possible to test it. Shall we try to measure variables or just observe? Which suits the research question better? Which would give us the most information *in this particular situation?* This leaves us without security blankets. We have to use our heads. It also means that only particular pieces of research are more or less scientific, depending on the theoretical importance of the questions, the judgment shown in the choice of methods, and the skill shown in applying them—as judged by our colleagues who do not always agree. Therein lies the risk.

References

Ford, J. *Paradigms and fairy tales,* Vol. 1. London and Boston: Routledge & Kegan Paul, 1975.

Glaser, B.G. and Strauss, A.L. *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine Publishing Co., 1967.

Kyburg, H.E., Jr., and Nagel, E. *Induction: Some current issues.* Middletown, Connecticut: Wesleyan University Press, 1963.

Kuhn, T.S. *The structure of scientific revolutions.* Chicago: University of Chicago Press, 1962.