theory approach, analysis need not remain at the substantive level. By taking the analysis to higher levels of abstraction and conceptual integration, grounded theory methods provide the means to develop formal theories (Glaser and Strauss 1971; Strauss 1978). To do so, the grounded theorist takes the comparative methods further. After developing conceptual categories, he or she refines and reworks the emerging theory by comparing concept with concept. Developing formal theories necessitates sampling a variety of different situational contexts and groups in which the concept applies. That way the theorist analyzes the boundaries and applications of the developing theoretical framework. To date, however, the grounded theory approach has been used primarily to develop rich substantive analyses. A theoretical analysis at the substantive level, though more modest in scope and power than formal theory, gives the analyst tools for explaining his or her data as well as tools for making predictions.

A Theory of Qualitative Methodology: The Social System of Analytic Fieldwork

Introduction

Readers of qualitative field studies repeatedly raise four questions about evidence. They may be characterized as four "R's" that haunt participant observers in sociology. I will illustrate these evidentiary questions as they might be addressed to a field study I conducted on civil lawyers for the poor. These illustrations will then provide a point of departure for a reanalysis of the methodological issues involved more generally in the assessment of fieldwork evidence.

1. Representativeness. I studied the legal assistance programs in Chicago through historical documents, organizational files, and eighteen months of fieldwork in 1972 and 1973. Since 1973, I have had numerous interviews with lawyers who have worked in Legal Services programs in Chicago, Connecticut, and California, but my report primarily covers experiences and events in Chicago in the early seventies. The study may have some value as a historical document of local interest, but I offer it as a general institutional analysis of the careers of legal assistance lawyers and their organizations. Can such generalization be justified?

2. Reactivity. In Chicago, I observed and talked with staff lawyers sometimes for minutes, sometimes for hours; in court, their offices, their homes; without anything approximating a fixed questionnaire. Perhaps the differences I report in the data simply reflect differences in my behavior. I might be asked, "How do you know it looks the way you describe it when you're not there looking?"

3. Reliability. I typed up about 2,000 pages of field and interview notes, presented only a fraction in the final report, and have not specified the criteria I used to select those data I published. A sympathetic but concerned reader might well observe and ask, "There is an infinite amount of background context that you could have included or excluded from your original field notes and the final text. The meaning of the behavior described would change with a change in the description of its context. How can you say your descriptions are the right ones?" A less kind

reader might put it, “How do we know you didn’t overlook disconfirming data, or even make it all up?”

4. Replicability. I began the study at a time when I learned that the legal assistance organizations in Chicago were about to merge. My initial interest was twofold, “Merger” seemed an attractively elusive phenomenon, one on which there seemed little useful sociological literature. The other interest was in the analysis of personal and collective careers and their relations. This interest grew out of my fascination with Georg Simmel’s writings, and from my training in symbolic interaction, with its perspective on the processes in which people shape individual and collective identities.

I began by interviewing organizational leaders, asking about their personal careers and their expectations for the merger. Then I began sitting in on lawyer-client interviews. When meetings began at the merged organization, I attended. I recorded my observations and interviews, sometimes contemporaneously, sometimes within an hour or two, often that night, occasionally weeks or even months later. The “career interviews” were loosely structured, to say the least. I would typically begin with questions about how the lawyer came to the job, move to initial experiences, then encourage a recollection of stages and changes in internal career, and finish by asking about expectations for the future. I changed the focus of my observations and interview questions in innumerable, unrecorded moves.

In light of such broadly formulated interests and inconstant methods, one might well ask, “If we wanted to test your analysis by repeating your research, how would we know what to do?”

Qualitative field studies appear especially vulnerable to criticism because they do not proceed from fixed designs. They do not use formats for sampling that could produce statistics on the representativeness of data. They abjure coding books that might enhance reliability by spelling out in advance the criteria for analyzing data. They fail to give detailed instructions for interviews — the questions to ask, their order, when to give cues and when to probe — that might give the reader faith that differences in subjects’ responses were not due to variations in the researcher’s behavior. Because of their emphasis on informal and flexible methods, qualitative field studies seem to make replication impossible.

Unfortunately, qualitative researchers have customarily conceded fundamental methodological weaknesses when faced with the four “Rs.”

1 “Qualitative” researchers have often called for a reorientation of methodological thinking, but they have not transcended the conventional approaches. Becker and others (1961:33–45) have proposed a post-hoc application, to informally gathered data, of methods used to guide the collection of data in more formalized research. For example, their remedy for reactivity is to count up all field notes bearing on a given proposition and to assign greater weight to observations of behavior in group displaying significant statistics, they acknowledge the “merely exploratory” status of their “case” studies. Citing the importance of “getting close to the data,” participant observation fieldworkers concede risks of reactivity (Scott 1965:266). Symbolic interactionists discard coding books, and along with them the goal of verification, as a necessary cost of developing “grounded” theory (Glaser and Strauss, 1967). As for replicability, some qualitative methodologists recommend reporting a natural history of the study but ultimately attribute success to “having the gift” (Lofland 1976:318).

Something important is missing from methodological thinking. On the one hand, as both Becker (1970:6–7) and Glaser and Strauss (1967:12) have noted, the sociological community, despite its neglect of the rationale for qualitative methods, frequently rewards qualitative empirical research. On the other hand, forceful questions have been raised as to whether formal research designs ever could be implemented according to their prescriptions. In fact, the argument increasingly goes, research designs never do anticipate fully the social relations that emerge in the research process. However random the sample, research subjects are never chosen through logical deduction from the theory purportedly tested (Camilleri 1962). Inevitably there are unscheduled influences on the understanding by respondents of the questions or stimuli (Cicourel 1964). Unexplicated bases for interpretation are ubiquitous (Garfinkel 1960), and they undermine the effort to establish that one has performed a replicable study.

These considerations raise two sets of related questions. Are not the methodologies of social surveys and experiments premised on a substantive sociological claim that supposedly pre-fixed features of the research, such as plans for probability sampling and coding rules, can actually determine and accurately predict the social relations established by the researcher with subjects, readers, and subsequent researchers? What is the methodological value of the formal research design if it does not in fact govern re-

settings than to reports of one-to-one encounters between researcher and research subject. They also recommend a post-hoc quantification procedure to create evidence on representativeness. These steps lead to the acceptance of a proposition about a group if the total number of observations and the ratio of positive-to-negative observations exceed arbitrary standards.

These are, at best, second-best solutions. When enumerated data are offered as the product of surveys and experiments, they purport to describe the precise number of instances underlying the substantive analysis, and they treat each datum as having a discernible weight on the analysis. The observational researcher who has directly entered, diffusely experienced, and variously recorded a natural setting cannot support this claim. Some field notes are, after all, based on observations covering seconds, others on a day’s experience. . . .

2 Despite the rhetoric of discovery and exploration, we are asked to attend to qualitative studies not merely on the claim that they develop attractive ideas but on an assertion that something “out there” has been discovered — on an empirical assertion that the theory is in fact “grounded.” For a vigorous critique of the distinction between discovery and verification, see Feyerabend (1975:165–169).
value in piling up data of a sort already determined to be consistent with the theory. Quantification therefore plays no logical role.³

I used analytic induction throughout my research on the careers of legal assistance lawyers. Legal assistance leaders had often complained about “high turnover” and staff lawyers had often remarked that “two years” represented a benchmark for assessing their careers. Was there, I wondered, a common process of leaving the institution, or “burning out,” as the lawyers put it? Was there a concise explanation of when and why staff lawyers would burn out?

My first step was to allocate into groups of short and long tenure all the lawyers who had entered the organization at least two years before the date of my analysis. In effect, the two-year point represented the initial definition of the thing to be explained. Then I looked for background features unique and common to those who remained more than two years, such as education, age, prior experience, political philosophy, work location, ethnicity, and sex. In effect I was looking for factors that would perfectly explain why lawyers were on one list or the other. This did not work, but instead of abandoning the effort I took an obviously artificial tack: I excluded all the confusing cases and drew up two neat lists for comparison, one with lawyers who had stayed more than two years and who had been “activists” before joining legal services, and another with lawyers who had left within two years and who had not been activists. Then I considered “exceptions” one by one, modifying the definition of the explanadum or the explanans in order to fit the “exception.” I was manipulating the meaning of the concepts distinguishing the lists in order to restore the perfect correlation which initially characterized them.

It quickly became apparent that I could not hope to explain the difference between those who did and did not stay more than two years. Some lawyers who had stayed more than two years were miserable, as unhappy as some who had left within a few months. Idiosyncratic factors, such as the chance appearance of job offers, might be what really distinguished the two. So I changed the definition of the explanadum to “desiring to stay two years.” This definition of the problem provoked new analytic difficulties. Legal Services programs offered staff lawyers a great variety of work set-

³Because of the irrelevance of quantification to the logic of analytic induction and because of the search for qualitative variation implicit in the hunt for negative cases, the following discussion frequently contrasts “qualitative” and “quantitative” research. In fact many, perhaps all, researchers use a combination of quantitative and qualitative methods. The possibilities for mutually beneficial combinations have been argued by Zelditch (1962), Reiss (1968), Sieber (1973), and Myers (1977), among others. The claim usually is that quantitative methods offer evidence or proof; qualitative methods, validity or insight. Whatever the merits of these conciliatory positions, they have failed to explain how qualitative methods can be rigorous in their own right.
necessarily associated with the method. As originally proposed, analytic induction claimed a superiority over “enumerative induction” by promising perfect correlations and “universal” explanations rather than probabilistic findings. But then very few if any perfect explanations appeared. Yet this embarrassment misconceives the methodology. Analytic induction ought to be evaluated in the same way in which field researchers practically gauge the value of their work. The test is not whether a final state of perfect explanation has been achieved but the distance that has been traveled over negative cases and through consequent qualifications from an initial state of knowledge. Analytic induction’s quest for perfect explanation, or “universals,” should be understood as a strategy for research rather than as the ultimate measure of the method. Analytic induction is a method for conducting social research, not a perspective from which to evaluate findings. . . .

In addition to their claim of “universal” explanation, early proponents of analytic induction (in particular, Znaniecki 1968) unnecessarily raised hackles by arguing its superiority over “statistical enumeration” for developing “genuinely causal laws.” Against this background, Turner’s critique—essentially that the concepts of explaining and explained phenomena in studies using analytic induction shade into each other and suggest tautology—was especially forceful. Analytic induction appeared to produce good definitions at best, not causal explanations.

The case for analytic induction can be made stronger with a number of revisions. If we view social life as a continuous symbolic process, we expect our concepts to have vague boundaries. If analytic induction follows the contours of experience, it will have ambiguous conceptual fringes. Its independent and dependent variables will inevitably shade into each other, suggesting tautology. But this weakness is only remarkable if exceptional claims are made for the method. Analytic induction and enumerative induction (in other words, survey statistics) differ in the form, not the fact, of uncertain results. For the statistical researcher, practical uncertainty is represented by statements of probabilistic relations; for the analyst of social process, by ambiguities when trying to code borderline cases into one or the other of the “explaining” or “explained” categories. In application to given cases, predictions on the basis of probabilistic explanations will sometimes be wrong, and predictions on the basis of explanations of social process sometimes so indeterminate as to be useless. (Turner did not claim that all explanations of analytic induction would be circular, nor that all its predictions would be indeterminate. See note 5.) . . .

A final difficulty for using the tradition of analytic induction is its apparent emphasis on an epistemology of “induction.” What field researchers actually do when they use analytic induction would be described more properly by philosophers of science as “retroduction” than as induction: a “double fitting” or alternating shaping of both observation and explana-

---

4 Since Turner’s (1953) review of the handful of studies then recognized as examples of analytic induction, there has been virtually no explicit discussion of the method. An exception is the extensive treatment in Lindesmith (1968) and Manning (1978); see also Moskos (1967:104–105). . . .
Theory and Evidence in Field Research

tion, rather than an ex post facto discovery of explanatory ideas (Hanson 1958:85ff; Baldamus 1972). To signal both my departure from several aspects of the tradition of analytic induction and my debt to the tradition’s essential guide to research practice — the injunction to search exclusively for negative cases — I drop the reference to “induction” in favor of the rubric analytic research.

Analytic field studies will not produce “proof,” i.e., artifacts of evidence which speak in a standard language or specialized fashion about representativeness, reliability, and so forth. The exclusive commitment to search for negative cases implies that there ought to be a different conceptual point for each reported phenomenon. Each datum reported should make its own substantive and not solely evidentiary contribution to the analysis. But analytic fieldwork does create an elaborate framework which can be used by researchers to assess how well they are doing and by readers to make evaluations. That framework is a social system, of which I sketch the following aspects.

Applied consistently in field research, the search for negative cases will: force the researcher to focus on social process as experienced from within; induce research subjects to act toward the researcher as a meaningful member of the native world; enfranchise readers as colleagues competent to make an independent analysis of the relation between data and explanation; and shape a role which subsequent researchers can readily take up for testing substantive findings. I suggest that this is a social system distinctively constructed by analytic fieldwork, in contrast with quantitative social research from fixed designs, and that this social system can be invoked to spell out answers to a wide variety of methodological questions frequently asked of qualitative field studies. This system of social research relations promotes generalizability, reduces the problem of reactivity, establishes constraints toward reliability, and enhances replicability.

Representativeness

The strategy of analytic research is to expand constantly the domain to which an explanation validly can be generalized. The sequential process in which theory is altered upon discovery of a negative case, in turn changing the meaning of a “negative case,” allows each new datum to function as a rival hypothesis. This method invests research energy with maximum efficiency to improve the generalizability of theory.

Analytic research rests the external validity of a study on its internal variety. The more differences discovered within the data, the greater the number of possible negative cases, and thus the more broadly valid the resulting theory. From a perspective on the sociology of social research, the analytic method, if it is followed, actually promotes the discovery of internal variety and thus its logic for establishing external validity.

A Theory of Qualitative Methodology: The System of Analytic Fieldwork

In practice, the analytic method shapes a particular researcher perspective on research subjects. It leads researchers inexorably to examine social process as subjects experience it from within. Once researchers have been led to examine the emergence in subjects’ experience of the phenomena to be explained, they find that their basis for generalization — qualitative variation — has expanded vastly.

This has in fact happened in every known instance in which analytic induction has been used expressly to discipline social research. At the start, the researcher’s conceptual units have often been static background factors and discrete acts. As the study has developed, the units have become processes with vague boundaries. Lindesmith (1968) discovered that he could not explain the first act of taking opiates, only addiction, a sustained use. Rejecting explanations by personality type, he offered a motive developed in the process of use, a “craving.” In my study, an early concern was to explain turnover among poverty lawyers. This was a focus on an act of leaving an organization. I ended with an explanation of “involvement,” a perspective on continuing a line of activity as intrinsically compelling...

If this hypothesis on the effects of the use of the analytic method on the researcher’s perspective is true, it implies a principle to guide qualitative research. Given the strategy of exploiting internal variety in order to warrant generalizability, the ideal site is one that is both in a period of historical change and has the most differentiated members. These do not appear to have been the principles typically guiding the selection of sites for qualitative research.

Critics might respond that the fact that a site is distinctive in the heterogeneity of its members and in the drama of its historical change makes it
unrepresentative. The researcher’s naturalistic focus on symbolic social process suggests a strategy to work on this problem. Take the charge that unique features of the research site — extreme differentiation of members, large scale of organization, rapid change in collective character — bias all the data collected. The qualitative researcher can examine the range and fluctuations of members’ situated experiences and may discover tests for the rival hypotheses. If the objection is that a smaller, more homogeneous and static context would alter members’ behavior, the researcher may locate exceptional members who, for a time at least, were situated in a homogeneous subunit isolated from the influences of a general historical trend.

To use this logic, one must assume that there are not complete discontinuities on the dimension at issue between the case studied and the place or time invoked in the rival hypothesis. For example, I tried to explain involvement and alienation from work among poverty lawyers by studying organizations in Chicago in the early 1970s. I would like to determine whether my theory requires qualification when applied to lawyers working in rural California poverty law offices, but I have no direct data. Is it reasonable to reject the hypothesis that the theory cannot be extended to the rural California site by examining the exceptional experiences of a lawyer who worked in Chicago’s Mexican-American neighborhood legal assistance office and cultivated vineyards in abandoned West Side lots in preparation for a move to a poverty law job in his native Northern California? I would also like to test the validity of my theory for lawyers currently assisting the poor. One might object that political commitment and its collective mobilization was unusually strong and pervasive among poverty lawyers during my fieldwork, in the early seventies, and that there is a general malaise now. But is the contrast so complete that a close examination of the experiences of the earlier group could not have encountered instances of the currently dominant perspective? By definition, no researcher can prove continuity between what he has and what he has not studied, but the analytic method points a way for thinking the problem through....

Quantitative and qualitative strategies toward generalizability strike different bargains with the existential limitations of social research. In attempts to establish statistical significance, the more the researcher sees data as heterogeneous (the greater the number of variables examined in a given number of data), the less likely it is that levels of significance will be reached. The goal of specifying the explanation by testing it against rival hypotheses through partial correlation or elaboration analysis may be pursued only at the cost of weakening the argument that the patterns examined

---

A Theory of Qualitative Methodology: The System of Analytic Fieldwork

are significantly representative of a larger population (see Camilleri 1962). Statistical evidence of representativeness depends on restricting a depiction of qualitative richness in the experience of the people studied. A similar practical trade-off confronts those who do inductive research, but it forces the opposite choice. By searching for data that differ in kind from instances previously recorded, analytic research creates a picture of the scene researched that is strategically biased toward much greater variation than random sampling would reveal. Brilliant qualitative studies such as Goffman’s Asylums (1961) overrepresent the richness of everyday life in the place actually observed in order better to represent social life outside of the research site.

Reactivity

...I have asserted as an empirical proposition (or perhaps more accurately, reasserted after Blumer [1969:82]) that when sociologists committed to hunt for negative cases examine theories that explain discrete acts by background psychological or social characteristics, they will inevitably transform their theoretical perspective into a focus on social process as experienced from within. When this research perspective is used in direct contact with subjects, it becomes a form of participant observation. Participant observation appears to exacerbate the problem of reactivity — subjects’ responses to a study’s methods that confound substantive findings. Among approaches to participant observation, the analytic method might appear the worst. ...Flexibility and fluctuation in research behavior are required. A participant observer committed to search exclusively for negative cases might constantly change the content of questions or the angle of observation; and as a result, any difference in the behavior of research subjects could be attributed to a change in the researcher’s behavior.

Spokesmen for participant observation have taken a defensive position on the issue, noting dangers for “objectivity” (Scott 1965:266) and for “contaminating” the scene examined (McCall 1969). Reasonably courageous sociologists might well be frightened off by such metaphorical warnings. To use participant observation appears to risk not only destruction of the scientific self but the pollution of society!

But interaction between variations in research methods and variations in members’ behavior does not necessarily produce a methodological problem. It does when the resulting behavior is irrelevant to the researcher’s objectives, or when the researcher fails to interpret correctly how he is perceived by members. Yet it is precisely on these grounds that analytic field research shows distinctive methodological strength. In contrast to research that attempts to fix the researcher’s behavior with a design for gathering data, the analytic field method makes valuable substantive data out of the responses of members to the researcher’s methods. Moreover,
A Theory of Qualitative Methodology: The System of Analytic Fieldwork

of access is ongoing, continuing from situation to situation and from the beginning to the end of each interview, in the researcher's efforts to establish and maintain rapport. Indeed, once the process of developing rapport is over and researchers with fixed questionnaires are ready to begin serious interviewing, qualitative researchers are often ready to leave. By this point they have realized which questions make no sense to an interviewee and have found substitutes that do. An appreciation of such qualitative distinctions is more important for the analytic researcher than learning which way the questions are answered this time.

Rich data are available in members' efforts to place a field researcher in a role and at a distance useful for native purposes. A process through which members attempt to keep the researcher further out is revealing of the nature of the scene studied. So are ploys by members to draw the researcher further in. On the former: For virtually the entire course of an eighteen-month field study, Wieder (1974) failed to build rapport with the residents of a halfway house for parolee-addicts. By examining his frustrations in "learning the code," and by investigating similarities in the alien roles residents shaped for him and for the staff, Wieder detailed the techniques used by residents for achieving segregation. The very fact that the residents persisted in reacting to the researcher as nothing more than an irrelevant researcher provided relevant data on the dominant culture in the institution. On the latter: Gusfield (1955) turned into data the sometimes frustrating reactions of Woman's Christian Temperance Union leaders to his efforts at maintaining a formal interview role. Cast by them not as an indifferent, neutral, scientific "researcher" but as an informed member of the public, he was berated and subjected to proselytizing efforts. The concerns of members about the boundary between outsiders and insiders and their ability to define it are significant features of all social systems.

Of course proffered interpretation of the meaning of members' behavior toward researchers may be wrong. But member behavior that has been shaped in response to the researcher's methods is not necessarily more problematic as substantive data than behavior shaped in any other interaction. Field researchers have missed this point. Common topics in the literature on participant observation concern whether members are lying, being superficial, or showing racial deference to the researcher. There is no fundamental difference between these problems of interpretation and those about whether members are lying, being superficial, or showing racial deference to each other.

Reliability

... Quantitative sociologists have developed complex measures of reliability, many of which have been described in the annual American Sociological Association publications of Sociological Methodology. One old
and relatively simple quantitative strategy for providing evidence of reliability suffices to indicate the apparently unreliable nature of qualitative field methods. If rules for coding are specified before data are gathered, the researcher can produce specialized, statistical evidence on the extent of agreement among "judges" who independently apply the scheme to the same data. This strategy is inconsistent with qualitative research. By definition, so long as a researcher's encounters with data are governed by preset coding rules, they cannot be exploited to develop qualifications in substantive analytic categories.

But qualitative research is not necessarily "impressionistic." The search for negative cases leads the qualitative researcher to a holistic analysis that binds propositions and data into an intricate network. Seen within the social relations analytic observers develop with members and readers, the network constrains the researcher toward consistency in selecting and interpreting data. Such a network holds together my thesis on the careers of lawyers for the poor.

I have argued that the social environment presented to legal assistance lawyers—clients, adverse parties, courts—characteristically defines the problems of the poor as insignificant. In turn, poor people's lawyers typically experience expectations that their work should be routine. To maintain intense involvement in client representation, the lawyers must struggle to treat problems as significant by doing a specific kind of work: reform. Because the environment calls for routine, their maintenance of involvement depends on reform.

An elaborate network of analysis and data underlies this summary statement. There are two main themes in the analysis, a warp and a woof, each of which has multiple strands. Thus "the environment" includes the expectations presented by clients and by the adversaries and court settings brought in their wake. The "reform" activities of the lawyers include not only litigation objectives but the creation of an everyday in-tradoffice culture that resists and transforms "routine" messages received from outside.

Each of these propositional strands is itself a combination of evidentiary threads. I support the assertion of a judicial expectation for routine treatment by direct evidence. For example, I quote a poverty lawyer's account of an instance in which a state court judge responded to his argument of a far-reaching constitutional issue literally by throwing the pleadings out of court. I treat some data as neutral on the proposition, for example, reports of courteous judicial hearings of routine motions. I offer many types of indirectly supporting evidence, for example, explanations by legal aid lawyers that a court's failure to comprehend routine arguments represents judicial senility or alcoholism or prejudice against the poor. A fortiori, judges experienced as having such incompetencies would appear to be unresponsive audiences for complex arguments. Similarly, varied evidence bears on the characterization of the expectations of clients and opposing counsel.

To convert disconfirming into confirming data, it was necessary to qualify concepts and generate explanatory propositions. On the generation of explanation: If the environment defines the lawyers' work as routine, then one should find that the lawyers' development of reform strategies is a necessary condition for their involvement in work. Further, if some lawyers who are litigating reform issues describe themselves as disengaged and demoralized, then another necessary condition must be added to the explanation. This second condition was found to be participating in a peer-sustained culture that expresses a reform perspective. On qualifying concepts: If lawyers who are not litigating for reform nevertheless recount extended periods of immersion in work, the theory must be refined by elaborating the meaning of "being in the institution's environment." I found that these lawyers were occupied with internal leadership projects of institution building such as training other lawyers, not with directly representing clients in an adversarial setting. The result: A complex analytic framework supports any proposition, although the framework is illustrated by what may seem superficially to be casually selected "anecdotes."

However unconvincing the reader may judge this institutional analysis of routine and reform to be, my point is to indicate the many ways in which it could be embarrassing. There is no insurance that analytic researchers will make rigorous interpretations, but readers can easily guard against being misled. As a result, as a practical matter the researcher faces strong constraints toward reliability. On the mundane level of mechanics, self-deception and biased selectivity in recording data will involve substantial difficulties.

Considering the social relations created in the research process, there are several methodologically salutary features of participant observing. Analytic field study builds relations such that the researcher will often be unable to grasp immediately whether what he is recording is supporting or contradicting his current analysis. I assume the following experiences with other qualitative fieldworkers. In the field I often wonder whether I should be elated or depressed for my theory in response to the course an interview or observation is taking. Group scenes usually contain

---

9 Holistic studies are usually thought of as case studies that try to comprehend an entire organizational or community social system. I am indebted to Diesing (1971) for his empirical research on the methodology of such studies. I believe that analytic induction takes on a holistic character even when it seeks to explain a particular line of action and that therefore the following methodological comments are applicable to analytic induction in general. Of his attempt to explain opiate addiction, Lindesmith (1947:15) wrote: "The actual process of the study may best be described as an analysis of a series of crucial cases which led to successive revisions of the guiding theory and to a broader and broader perception of the implications of that theory. Isolated bits of information and apparent paradoxes one after the other seemed to form integral parts of a consistent whole."
much that is obvious to members but challenging for me to comprehend. In interviews I must restrain analytic commentary in order to remain respectfully attentive and in order to provoke respondents to keep responding. I have no forms on which observations can be checked off and no set formulas for probes. A fieldworker inclined to ignore disconfirming data and record only confirming data often could not easily make the discrimination.

Once the qualitative researcher is out of the field and constructing a text, the social relationship of writer to reader presents elaborate constraints against inconsistent and unexplained interpretations. If the qualitative data-gathering and text-construction process seems inarticulate, even mysterious, it helps to recognize that, irrespective of how unruly the analytic researcher’s practices, the reader has rules available to detect error in the text. Blumer’s classic critique of The Polish Peasant (1939) demonstrated the multiple objections that a discerning reader could make to qualitative research reports. Charges of a lack of fit between data and analysis may come from many sources: from multiple interpretations by the reader of the data presented; from the apparent irrelevance of the member’s meaning to the analyst’s point; from the connection between the analysis and the data being made through interpretive commentary rather than through the data itself; and from inconsistent implications of data presented in different parts of the text.

The weblike character of the text means that each datum will ramify in implications throughout. To insulate a careless analysis from critical readers, the researcher would have to engage in a laborious process. Each quote or episode would have to be edited carefully so that it might avoid contradiction elsewhere in the analytic framework. For example, if I had characterized the legal assistance lawyer’s professional environment loosely as disreputable or demeaning (one of my earliest hypotheses), then I would have had to purge, from all quotes, any indication that a local judge or opposing counsel may have acted respectfully. To protect the initial, casual analysis, an extensive chopping up of quotes would have been essential, and further, a meticulous effort would have been necessary to avoid the appearance of chopped data.

Authors of qualitative field research reports cannot escape a dialectical evidentiary bind. The analysis must be made dense to make the data representative, to claim, in other words, that the study is generally useful. If the network of field materials and propositions is very limited, it would be easy to indulge inclinations not to report inconsistent data. Of course this could be done, but after a point the deceit would become self-defeating. Who would care? The study then would not pretend to be very useful or significant.

Given the possibilities of misfit, a biased selection of data that would convince a careful reader is not easily achieved. Given the emergent character of the analysis, if a confirming quote is hard to find or invent, the alternative readily available is to alter the analysis so that the data at hand will suffice. The everyday stuff of writing qualitative analysis consists of an ongoing series of retrodictive shifts: trying to convince oneself that a quote or episode can be interpreted to fit the analysis until frustration is sufficient to make stepping back and modifying the analysis seem the easier course.

In the traditional view, qualitative fieldworkers seem relatively free of practical constraints on recording and interpreting data wishfully and carelessly. Analytic research must be kept small scale in its human organization. Arrangements to deploy numerous researchers and coordinate their activities would compromise the method by requiring a prespecification of the data that they are to look for. Thus little if any mutual consent must be achieved to invent qualitative field notes. Qualitative research produces bulky field notes recorded with abbreviations meaningful only to the researcher. The interpretation of field notes often depends on a knowledge of context supplied by prior field notes or known but not recorded by the researcher. Field notes cannot as readily be transferred to other sociologists in original state as can responses to fixed-choice questionnaires because they cannot as easily be masked to preserve confidentiality without altering their meaning. The analytic strategy, which never separates data gathering from inspection of “results,” may tease the qualitative researcher to disregard disconfirming data selectively, perhaps through an unconsciously biased inability to understand “inarticulate” responses. In contrast, the collection of quantitative data from a preset design may block the researcher’s awareness of what findings would be disconfirming until data collection is complete and the computer has finished its run. The rules which preset the meaning of data to be gathered through surveys are used in large-scale research as a framework for an organizational hierarchy which gets the work done. An elaborate conspiracy might be necessary to manufacture findings. Moreover, the frequent practice in survey research of hiring specialized data gatherers who lack responsibility for analysis would appear to insure motivational neutrality.

On the other hand, this arrangement carries the risk of building alienation and indifference into a study at its most basic level (Roth 1966). In contrast, the close relationship between field researcher and subjects should make it more likely that the researcher will take the people studied as significant others. This audience can provide powerful constraints on reliability. To dismiss their objections to interpretations, the researcher might have to renounce an emotionally significant segment of his or her life. It would also seem to be easier to alter the number in a category than to invent quotations that sound like seventy-five different research subjects. Working with hypotheses, one could specify statistics on significance and correlations which would be confirming. One could instruct a machine to
Theory and Evidence in Field Research

figure elegant equations backwards and manufacture the data necessary to make the math succeed. Just as it would be easier to change the number entered in a category than to invent a quote, it would be easier to figure out what that number should be.

My purpose is not to impugn the integrity of statistical researchers but to outline an empirical theory for evaluating reliability in analytic field studies. I have used the issue of manufacturing data as a way of short-cutting a more lengthy argument that would cover in detail allegations of morally lesser methodological sins. If there are constraints inherent in analytic field research which automatically place the dishonorable researcher between the Scylla of apparent unreliability and the Charybdis of apparent insignificance, a fortiori the merely careless analytic researcher should be found in the same straits. To develop in detail a theory of the constraints against fraud in qualitative and quantitative social research, one might examine real cases of serious allegations. But for the present, if the methodological strength of research depends on the social system it actually fashions, qualitative field researchers need not be deferential in evidentiary debate.

Replicability

To the extent that researchers pre-fix their decisions for gathering data, they can easily present readers with a format for testing findings by repeating the study. Questionnaires and sampling procedures defining the boundaries of the relevant population may be included in an appendix; the coding book and written instructions for administering the survey instrument may be copied and mailed to subsequent researchers. Apparently inviting replication, psychology experiments traditionally have been reported in articles that neatly separate the description of methods and findings. The format takes the posture: You don't believe it? Go see yourself.

Analytic field research changes procedures for gathering data in order to encounter negative cases, then changes the analysis, and so on, in an interactive relation of method and substance. Innumerable ad hoc judgments are made in the field, decisions on when to visit the research site and when to move from observing one situation to another, decisions on when and how to probe in interviews. They could be reported, if detected, only through retrospective reconstruction. Because standards of substantive relevance change rapidly within the research process, much of the data considered will not be reported nor even recorded. The difficulties of specifying the research procedures used and of accounting for all the data considered add up to an inability to define what a replication would be.

Despite these facts, the analytic field research strategy promotes relations with other researchers that facilitate the subsequent testing of substantive findings. . . . [T]he claim in analytic research that no negative case can be found invites the testing of findings without repeating the original research. A subsequent researcher can simply pick up where the study left off, looking for a single contradiction.

If the costs of subsequently testing qualitative field research findings are relatively low, so the rewards are relatively high. It has been notoriously difficult to publish failures to reject null hypotheses. Publishing criteria are biased toward disconfirming and innovative results. An attempt to replicate a study with a fixed design and determinate findings runs the risk of becoming nothing more than an unpublished confirmation. The risk is a significant deterrent. In contrast, subsequent tests of qualitative field studies will never be merely attempts at disconfirmation. If only because the original research fails to specify what an exact repetition would be, a subsequent researcher should be confident of documenting new types of phenomena, valuable for other theoretical purposes, in his search for disconfirming cases.

Analytic field research also more democratically empowers readers to become subsequent testers. . . . Qualitative research reports properly may be regarded as good to the extent that readers test them in application to new data in the very process of reading. Underlying the reader's experience in "recognizing" as valid or rejecting as "artificial" an analytic formulation in a qualitative text is an implicit application to phenomena within the reader's experience, to new data existing beyond the reach of the original research.

To appreciate the reality of such testing, compare the implications of two allegations, that when writing Asylums, Goffman invented his portrait of the mental institution; and that Hollingshead and Redlich invented the survey responses and the computations presented in Social Class and Mental Illness (1958). Assume it is 1961, and follow-up studies have not yet been attempted. Readers of both works would, I submit, respond differently. There is a sense in which a reader would judge that the former charge could not be true. If Goffman was never a participant-observer in St. Elizabeth's Hospital, as he said he was, he must have been in some other mental hospital; or he must have talked to people who were; or read accounts by people who were. Ignorant of his methodology, one takes his results as evidence that he did something right. One can judge the value of his text immediately with as wide a variety of methods as he might have used. For the quantified survey study, the allegation of dishonestly reported methods and fabricated findings is much more crucial. One can readily imagine how the allegation could be true; if one wants to test it, one faces a sizable task; and if one believes the allegation, the work is worthless. To evaluate such claims requires an accurate and detailed account of how the findings were produced from a pre-fixed design.

10 In an informal note, Mel Pollner suggests to me that the inquiry might start with a comparison of the controversies around the work of anthropologist Carlos Castaneda (Strachan 1979) and psychologit Cyril Burt (Hearmshaw 1979).
In a fundamental way, the allegation of fictive data is less meaningful when applied to qualitative field studies. In fact, many of the best interweave observational and interview data with excerpts from novels written by earlier participant observers. An example is the use of Melville’s *White Jacket* in Goffman’s *Asylums* (1961:33–34). Another legitimate use of fictive data is illustrated in a book by Rosett and Cressey (1976). Drawing on wide but unspecified prior research and participant observation experience, they invent a criminal case, a cast of players, and a multistage decisional process—a whole social organizational setting and drama—in order to demonstrate the collective construction of guilty pleas. . . . When such authors blur the line between fiction and data in their texts, they are obeying tendencies natural to their methodology. Phenomenologically, the distinction between “created” and “recalled” data becomes ambiguous in qualitative field research. Observations can be recorded at any time, contemporaneously or long after they are made. No rules govern the timing. Researchers can credit as data their own experiences in interaction with members.11 Given this methodologically sanctioned freedom, the researcher may often be unable to assert confidently whether his image of a research scene is recalled or “made up.”

But given the relation between author and subsequent researcher, this is a very constraining freedom. Qualitative researchers obtain no license from their affinity to the novelist. After all, the analogy between the novel and participant observer’s qualitative text is not complete. The requirement for an explicit analysis that is more general than the case under study, plus the discipline of the negative case, breed a compelling concern that one is not manufacturing data. On what else of the accuracy of his analysis in the scene researched can the author rely to avoid a subsequent researcher’s discovery of disconfirming data and the consequent charge that the analysis offered exists only in its author’s mind? For the analytic researcher, methodological constraints are experienced as existential matters, not as matters of methodological convention.12

Researchers’ Social Relations and the Evaluation of Analytic Fieldwork

By recommending a sociological perspective on methodology, I mean to call attention to implications for the evaluation of findings that may be discovered by examining the system of social relations created in the research process. I have proposed hypotheses on the social relations created by analytic field research with reference to four familiar standards of evaluation. First, representativeness. By searching exclusively for negative cases, a researcher gives distinctive shape to his or her perspective on the people studied. To avoid contradiction by negative cases, analysis inevitably turns to the examination of social process as experienced from within. Continued study of a given segment of social life turns up increasingly refined discriminations between states of phenomena as they emerge in and vanishes from experience. . . .

Researchers who gather data from fixed designs acknowledge that statistics on significance measure only the uncertainties of extrapolating from the data examined to the specific population sampled, not the prospects for extrapolating to later times and other places. Since any use of the findings of a study will necessarily be in application to a different population, a showing of statistical representativeness is necessarily an incomplete achievement. Analytic research cannot measure the uncertainty with which the scene it describes reflects any larger population, but its single-minded pursuit of qualifications in order to enhance the prospects of accurate application to other times and places raises the question whether this falling should be troublesome.

In its present state, the methodological literature assumes that reactivity in participant observation is a contaminating problem. But if we examine how research procedures shape the meaning of the study to members, we may conclude that field research without a formal design makes interaction between researcher and member into a substantive data resource. In a sense, analytic field research dissolves the problem of reactivity, whereas formally rigorous methods actually create it by enhancing the appearance of “research” and by limiting variation in the meaning of the research process to members.

Qualitative researchers typically concede an inability to verify the reliability of their interpretations of data. Yet even sociologists who labor to show high levels of statistical agreement among indicators acknowledge a logical gap between indicators and what is indicated (Blacker 1968). True, field researchers make themselves vulnerable in special ways to questions of objective interpretation by elaborating a network of idiosyncratic observations and informally fabricated explanations. But does this imply methodological weakness? From a sociological perspective on the relation between researcher and reader, the analytic method confers on readers unique powers to make their own judgments on reliability from independent encounters with data. Is that not preferable? Unlike statistical measures of the relations between several items in an index, or of agreement scores on a given item administered at different times, the analytic qualitative approach does not separate the evaluation of reliability, or con-

---

11 See Glaser and Strauss (1967:252) for a defense of one such instance: Fred Davis’s article “The Cabdriver and His Fare” (1959), which was based on personal experience but written long after Davis left cab driving and without benefit of contemporaneous field notes.

12 Compare the distinction drawn by W. James (1970) between truth as “reflection” and a “pragmatic” theory of truth.
sistency in interpretation, from the evaluation of validity, or the mesh
between the researcher's concepts and the meanings expressed by sub-
jects. Is that not preferable?

Qualitative researchers often admit to an inability to describe what
would have to be done to repeat a study. They also have questioned
whether a study is ever so disciplined by predefined designs that, by adopt-
ing a published account of social research, anyone could ever really repeat
it. The strength of analytic field research for replicability lies in the rela-
tionship it creates between original and subsequent researchers. When used
in participant observation, the analytic method induces the researcher to
credit as data his or her own experience as a member, minimizing the
barriers to subsequent researchers for continuing the verification process.
If qualitative field research offers no insurance that the researcher did not
make up findings, it also raises the question of whether social research
need be conducted in such a way that fabricated data can become a mean-
ingful problem.

The analytic approach to fieldwork maintains an interaction between
method and substance, breaking down their separation. Analytic field
studies will not produce "proof," or artifacts which stand apart from sub-
stantive findings and can measure, or otherwise speak in a standard
language about, representativeness, reliability, and so forth. The process
and perspective for evaluation must be different. Acting within a system
of social relations constructed in the research process, readers can make
evaluations and researchers can assess how well they are doing.

Applied consistently in field research, the search for negative cases will
force the researcher to focus on social process as experienced from within;
induce research subjects to act toward the researcher as a meaningful
member of the native world; enfranchise readers as colleagues competent
to make independent assessments of the trustworthiness and general sig-
nificance of the analysis; and facilitate subsequent tests of the findings
by readers. This social system, which contains dimensions that have been
barely outlined in this preliminary essay, can be invoked to spell out an-
swers to numerous methodological questions frequently asked of qualita-
tive field studies. For each qualitative field report, readers can assess how
richly the researcher has perceived internal variation in the data; how
radically the researcher varied his approaches to subjects; the density into
which data and analysis have been interwoven; and the practical ease of
testing the theoretical claims on new data.
Contemporary Field Research
A Collection of Readings

Robert M. Emerson
University of California, Los Angeles

Little, Brown and Company
Boston   Toronto