
ERIC/ECTJ Annual Review Paper: Criteria for Assessing the Trustworthiness of Naturalistic Inquiries

Author(s): Egon G. Guba

Reviewed work(s):

Source: *Educational Communication and Technology*, Vol. 29, No. 2 (Summer, 1981), pp. 75-91

Published by: [Springer](#)

Stable URL: <http://www.jstor.org/stable/30219811>

Accessed: 20/12/2012 10:12

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Springer is collaborating with JSTOR to digitize, preserve and extend access to *Educational Communication and Technology*.

<http://www.jstor.org>

ERIC/ECTJ Annual Review Paper

Criteria for Assessing the Trustworthiness of Naturalistic Inquiries

Egon G. Guba

Egon G. Guba is professor of education, Smith Research Center, Indiana University, Bloomington, IN 47405.

This is the ninth ERIC/ECTJ Annual Review Paper, preparation of which was supported by the ERIC (Educational Resources Information Center) Clearinghouse on Information Resources, Syracuse University, Syracuse, New York. The material in this article was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Education. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not necessarily represent the official view or opinion of NIE. The author is grateful to David Clark, Larry Havlicek, Robert Heinich, John McLaughlin, Cecil Miskel, and Robert Wolf for their careful critique of an earlier version of this paper, and to his wife, Yvonna Lincoln, from whose joint work with the author in other contexts many of the ideas expressed in this paper emanated.

ECTJ, VOL. 29, NO. 2, PAGES 75-91
ISSN 0148-5806

This paper deals with the topic of criteria for judging the trustworthiness of inquiries conducted within the naturalistic inquiry paradigm. That paradigm, also referred to as the phenomenological, anthropological, or ethnographic, has become increasingly popular in recent years; reports of investigations carried out in this mode often take the form of case studies or "portrayals" (Stake, 1975) and bear little resemblance to the kinds of articles we have been accustomed to seeing in our leading professional journals. Both the novelty of the paradigm and the strangeness of the reporting format pose special problems for editors and referees of journals, peer review committees or dissertation committees considering proposals, and naturalistic investigators themselves as they attempt to design and monitor their inquiries. What precisely are the criteria that ought to be applied to this class of investigations?¹

¹This paper will focus on only a subset of criteria, those dealing with the trustworthiness of findings. There are other methodological problems that are special to naturalistic inquiry that cannot be dealt with here. Among these are: bounding problems, that is, issues dealing with the means by which inquirers determine what to include and exclude from consideration; focusing problems, that is, issues dealing with the means whereby inquirers organize and ascribe meaning to the data they do collect; and the problems of investigator competence. The last-mentioned is especially important since, as will be seen, inquirers often act as instruments (problems of the human as instrument). Hence it is important to know about the inquirer's training and experience in reaching a judgment about the trustworthiness of his or her data. This consideration has been omitted from the present paper since it has been assumed that the refereeing of journal

What follows is a primitive effort to answer that question. The effort is described as "primitive" because, while practitioners of naturalistic approaches have been reasonably introspective about what they do, they have not made systematic efforts to codify the safeguards that they intuitively build into their inquiries. Hence there is little prior conceptualization on which to build. It is my hope that this article, while it cannot be definitive, will at least open discussion on the matter and will lead, in due course, to better formulations than are possible now.

The paper consists of five parts. First, naturalistic inquiry is defined and differentiated from the rationalistic² mode of inquiry commonly practiced in the past and still dominating today. Second, the issue of *what* aspects of trustworthiness should be addressed by the criteria is raised, and it is proposed that there are four: internal validity, external validity, reliability, and objectivity (although these terms are translated into their naturalistic

counterparts). Third, a brief explication of how these questions are dealt with in conventional inquiry is presented. Fourth, a parallel or analogous mode for dealing with them within the naturalistic paradigm is outlined. Finally, implications of the proposed criteria are described. The exposition of the paper will be necessarily brief because of limitations both on space and in the current state of the art.

WHAT IS NATURALISTIC INQUIRY?

It is beyond the scope of this paper to provide a detailed description of naturalistic inquiry.³ Interested readers are referred to Guba (1978, 1979), Guba and Lincoln (in press), Wolf (1979), and Wolf and Tymitz (1977). The reader should be aware, however, of the following:

The term "naturalistic" describes a paradigm for inquiry, not a method. There are many paradigms for arriving at "truth," including, for example, the legal paradigm that characterizes courtroom proceedings; the "expert judgment" paradigm that characterizes accreditation site visits, peer review of proposals, and judgments of athletic performance; and the modus operandi paradigm (Scriven, 1976; tracing of characteristic causal chains) that characterizes forensic pathology and television repair.

Chief among the paradigms that have been utilized in support of disciplined inquiry are the rationalistic and the naturalistic. There is no basis for choosing one of these paradigms over others in each and every inquiry situation. Rather, each rests on certain assumptions that must be tested in the context of application. Just as it is proper to select that analytic statistic whose assumptions are best met by a set of data, so is it proper to select that paradigm whose assumptions are best met by the phenomenon being investigated. Particu-

papers will continue to be managed by a "blind" method; that is, that the author's name and credentials will not be among the materials that the referee is asked to take into account. All three problem areas are extensively discussed in Guba and Lincoln (in press); the former two in Guba (1978).

²In previous discussions (Guba, 1978, 1979; Guba & Lincoln, in press) we have referred to what is called here the rationalistic paradigm as either the *scientistic* or the *scientific*. The use of even the less pejorative of these latter two terms now seems to us inappropriate on two counts. First, readers have tended to view the naturalistic paradigm as less scientific (or even as nonscientific) and have, therefore, because of the enormous legitimation accorded to anything scientific in our culture, denigrated it as "less valid." Second, several critics have accused us of setting up a straw man, on the grounds that vanguard scientific thinkers have moved beyond the 19th century epistemology of logical positivism that we describe and ascribe to inquirers today. It is undoubtedly true that many scientists now think differently, but such change does not characterize the large majority of "scientists" who are engaged in inquiries today in either the "hard" or the "soft" sciences. There, the old culture still dominates. It is to that level of practice that our criticisms are directed, and it is of that moribund culture that our descriptions are apt. However, to avoid the unintended meanings that some readers have drawn from our work, we have shifted to the term *rationalistic* to describe the paradigm that informs conventional inquiry.

³Indeed, it would not be possible to provide such a description at this time. Especially lacking are good examples of naturalistic studies that have been completed and reported. ERIC can provide a valuable and needed service by establishing a data bank of such studies.

lar methods can be used in support of either (or any) paradigm (see comments on postures, below).

The naturalistic and rationalistic paradigms differ on certain key assumptions. These include:⁴

- *The nature of reality.* The rationalistic paradigm rests on the assumption that there is a single reality upon which inquiry can converge, and that that reality is separable or fragmentable into independently manipulatable parts (commonly called variables). Thus, certain variables can be singled out for study (or control) without essentially influencing others. The naturalistic paradigm rests on the assumption that there are multiple realities, that inquiry will diverge rather than converge as more and more is known, and that all "parts" of reality are interrelated so that the study of any one part necessarily influences all other parts.

- *The nature of the inquirer/object relationship.* The rationalistic paradigm rests on the assumption that the inquirer can maintain a discrete (and discreet) distance from the objects of the inquiry; that is, that the relationship between the inquirer and the object is essentially one of independence. The naturalistic paradigm asserts, instead, that the inquirer and the respondent (note the shift in terminology from "object") are interrelated, with each influencing the other. Of course naturalistic inquirers make every effort to maintain an optimal distance between themselves and the phenomenon, but never for a moment do they consider that the "optimal" distance is impervious to inquirer-respondent interchanges.

- *The nature of "truth statements."* The rationalistic paradigm rests on the assumption that generalizations—enduring truth statements that are context-free—are possible; indeed, it is frequently asserted that inquiry would have no point if this were not true. The rationalistic approach

aims at developing nomothetic knowledge and hence focuses on the *similarities* between objects of inquiry (similarities being the stuff out of which generalizations are made). The naturalistic paradigm rests on the assumption that generalizations are not possible, that at best what one can hope for are "working hypotheses" that relate to a particular context. The naturalistic approach aims at developing idiographic knowledge, focusing on *differences* between objects as frequently and with as much interest as on similarities.⁵

The issue here is *not* which assumptions are "true" but which offer the best fit to the phenomenon under study.⁶ More and more investigators have become convinced of the relative utility of the naturalistic paradigm for studying that class of phenomena that often is called social/behavioral. Social/behavioral phenomena exist chiefly in the minds of people, and there are as many realities as persons. Reality in this sense, moreover, gives the appearance of "whole cloth"; if one attempts

⁵It is of interest that while most inquiry develops from a nomothetic posture, applications must always be made in some idiographic setting. There is thus a major value mismatch in fields like counseling and guidance, school administration, and reading in which the research, done in the rationalistic mode, finds no application in individual cases (clients, schools, non-readers, and so on). The situation is similar to that confronted by particle physicists as formalized in the Heisenberg Uncertainty Principle. Statistical laws have no individual application.

⁶Assumptions are definitively not "self-evident truths" as the axioms of Euclidian geometry were represented to us in high school mathematics, but are, instead, statements taken for granted by agreement (or for the sake of the game). Whether the findings that result from the application of a particular inquiry have anything to do with the phenomenological world is a matter for empirical test, not argumentation. Thus, while Euclidian geometry is splendid for dealing with earth-sized phenomena, Lobachevskian geometry, based on different axioms, works very much better with astronomical phenomena. The question of which assumptions are "right" is irrelevant; rather one uses that geometry whose theorems are found to be valid for the space being dealt with. Similarly, we should not ask whether the assumptions of the rationalistic paradigm are more or less "right" than those of the naturalistic, but rather which paradigm provides a better "fit" to the phenomenon we seek to understand.

⁴The three assumptions presented here are those most salient for the analysis to follow. Other assumptions also differentiate the paradigms. The most interesting among these others is the extraordinary value claim of rationalists that their inquiry is value free.

to focus attention on certain portions of reality, the whole falls apart as though the cloth had been cut with scissors. Further, while investigators may be able to maintain a neutral posture with respect to physical or chemical phenomena (although even that possibility is called into question by the Heisenberg Uncertainty Principle), it is impossible to do so when the objects of investigation are people. Indeed, not only do respondents exhibit reactivity but so also does the investigator! Finally, human behavior is rarely if ever context-free; hence knowledge of human behavior individually or in social groups is necessarily idiographic, and differences are at least as important as similarities to an understanding of what is happening. To the extent that the conditions described are so, the naturalistic paradigm becomes the paradigm of choice.

The rationalistic and naturalistic paradigms differ in terms of certain postures which characterize their practitioners. While not compelled to do so by the logic of the paradigms they follow, practitioners of the naturalistic and rationalistic paradigms have displayed a propensity to take opposing positions along certain key dimensions. While other variations in posture might be described, the following afford some insight into the characteristic differences in approach taken:

- *Methods.* Rationalistic practitioners have preferred quantitative methods while naturalistic practitioners have preferred qualitative methods. This predisposition is so intense that the conflict between the two paradigms has frequently been mistaken for a conflict between quantitative and qualitative methods, a mistake in logic that has led to the generation of a great deal more heat than light. But of course these two dimensions are orthogonal; there is no inherent reason why either paradigm cannot accommodate, and be contributed to, by either methodology.

- *Quality criterion.* Proponents of the rationalistic approach have insisted that the single most important criterion for assessing the quality of an inquiry is its rigor, while proponents of the naturalistic approach argue for relevance. The distinction has sometimes been characterized as the

difference between saying, "It doesn't matter what you do so long as you do it well," versus saying, "Anything not worth doing at all is certainly not worth doing well!" These criteria are both worthy of attention, but unfortunately they are in a trade-off situation: The more one insists on rigor (internal validity), and assures it by control of the sort possible in a laboratory, the less relevance (external validity) one can expect, for the results will apply only in another laboratory.

- *Source of theory.* Adherents of the rationalistic paradigm prefer a priori theory, usually of the hypothetico-deductive type; indeed, such theory is indispensable since the rationalistic approach requires the statement of hypotheses to be tested or questions to be answered in advance. Such hypotheses or questions can be generated only from theory existing before the fact. Adherents of the naturalistic paradigm prefer to have the theory emerge from the data themselves, that is, they wish the theory to be grounded (Glaser & Strauss, 1967) and typically prefer their theory to be of the pattern type (Kaplan, 1964).

- *Knowledge types used.* Polanyi (1958) distinguishes propositional knowledge—knowledge that can be cast into language form—from tacit knowledge—knowledge such as intuitions, apprehensions, or feelings that cannot be stated in the form of language but are somehow "known." Everyone "knows" more than he can communicate, even to himself. Rationalistic inquirers operate solely at the level of propositional knowledge, primarily because everything they investigate is cast into the form of hypotheses or questions based on a priori theory. Naturalistic inquirers insist on the opportunity to build upon and expand their tacit knowledge as well.

- *Instruments.* Rationalistic practitioners typically interpolate a "layer of instrumentation" between themselves and the phenomena to be studied, partly because it is believed that by thus removing themselves from direct contact they will improve the reliability and objectivity of the study, and partly because it is felt that such "objective" instruments can be sharpened and refined to a greater level of sen-

sitivity than can a human observer (both assumptions are questionable; see Guba & Lincoln, in press). Naturalistic inquirers, on the other hand, are inclined to use themselves as the instruments, willingly trading off some objectivity and reliability (in the rationalistic sense) in order to gain greater flexibility and the opportunity to build upon tacit knowledge (a feature that paper-and-pencil or physical instruments can never have).

- *Design.* Rationalistic investigators insist on a *preordinate* design (Stake, 1975), that is, they require that every step from problem specification through data collection and analysis to reporting be described in advance. Of course their dependence on a priori theory, propositional knowledge, rigor as a quality criterion, and nonhuman instrumentation (which must be developed) more or less requires that such a posture be taken. Moreover, the typical experimental design used by rationalists prohibits any change in treatment once the study is under way lest the variances be confounded, disallowing interpretation of the results. Naturalists, on the other hand, believing in unfolding multiple realities, in interactions with respondents that will change both the investigators and the respondents over time, and in grounded theory, will insist on an *emergent* (unfolding, rolling, cascading) design, which is never complete until the inquiry is arbitrarily terminated as time, resources, or other logistical considerations may dictate.

- *Setting.* The adherent of the rationalistic paradigm prefers to conduct inquiries in the laboratory, since the laboratory affords the epitome of control. The naturalist, on the other hand, prefers to conduct inquiries in nature, inviting whatever interference the real world can provide. It is as though the rationalist is interested in what happens in the best of all possible worlds, while the naturalist is concerned with what happens in the worst.

While adherence to one or another of these sets of postures is not required nor logically compelled by the underlying axioms, followers of the paradigms seem to be strongly inclined to do so, probably because they were so trained. The assumption of one of these postures has become

identified with the proper way to do research—a kind of orthodoxy. But such intransigence is unfortunate. Once the decision to use a particular paradigm has been made on grounds of “best fit” of assumptions, compromises on postures are not only possible but well advised, regardless of which paradigm has been selected. Thus, both quantitative and qualitative methods should be used as the situation warrants. To seek an appropriate balance between rigor and relevance seems sensible. A priori theory can be grounded through earlier inquiry, and to insist that there should have been grounding at some prior stage does not seem unreasonable. Both tacit and propositional knowledge are useful, and, indeed, it is probably the hallmark of competent investigators that they translate tacit knowledge into propositional knowledge as quickly as possible. Astute investigators will utilize both themselves and other instruments, depending on circumstances. Some elements of design can always be specified in advance, and the wise inquirer will specify all such possible elements while retaining a flexible posture that permits changes and emendations as the situation may dictate. Finally, information from both the laboratory and the real world has utility in achieving understanding. Nevertheless, it seems clear that both camps have evolved orthodoxies, and *that* is a fact of life that must be taken into account in specifying criteria for judging trustworthiness that will be found generally acceptable.

WHAT QUESTIONS OF TRUSTWORTHINESS SHOULD BE ADDRESSED?

Guba and Lincoln (in press) suggest that four major concerns relating to trustworthiness have evolved, and it is to these concerns that the criteria must speak. The concerns as described by these authors are these:

1. *Truth value.* How can one establish confidence in the “truth” of the findings of a particular inquiry for the subjects (respondents) with which and the context in which the inquiry was carried out?

2. *Applicability.* How can one determine the degree to which the findings of a par-

ticular inquiry may have applicability in other contexts or with other subjects (respondents)?

3. *Consistency*. How can one determine whether the findings of an inquiry would be *consistently* repeated if the inquiry were replicated with the same (or similar) subjects (respondents) in the same (or similar) context?

4. *Neutrality*. How can one establish the degree to which the findings of an inquiry are a function solely of subjects (respondents) and conditions of the inquiry and not of the biases, motivations, interests, perspectives, and so on of the inquirer?

These four terms typically have been labeled within the rationalistic paradigm, as shown in Table 1; the parallel naturalistic terms are also shown. The "translation" requires some justification.

Truth Value. Within the rationalistic paradigm, internal validity is logically determinable by demonstrating isomorphism or verisimilitude between the data of an inquiry and the phenomena those data represent—not an unreasonable expectation when one begins with an assumption of a single reality upon which inquiry can converge. But it is not possible to test isomorphism directly—to do so would require absolute knowledge of what the real world is like. Instead, rationalists fall back on the strategy of ruling out all plausible alternative explanations. Thus, one consequence of the rationalistic approach is that hypotheses can never be directly confirmed (since a test of isomorphism is not possible) but they can be disconfirmed (by

showing that a plausible alternative hypothesis has a high probability of being correct). Nevertheless, the fundamental idea of isomorphism is useful, for within the naturalist's framework, the analog of isomorphism to reality must be isomorphism to respondents' perceptions (multiple realities existing in the minds of people). In establishing truth value, then, naturalistic inquirers are most concerned with testing the credibility of their findings and interpretations with the various sources (audiences or groups) from which data were drawn. The testing of credibility is often referred to as doing "member checks," that is, testing the data with members of the relevant human data source groups.

Applicability. Within the rationalistic paradigm, applicability—external validity or generalizability—requires that the inquiry be conducted in ways that make chronological and situational variations irrelevant to the findings. If that condition can be met, the findings obviously will have relevance in any context. Generalizations are taken as enduring; that is, unchanging over time, truth statements that are context-free—that hold in *any* context. But Cronbach (1975) has argued that all generalizations "decay" like radioactive substances, having half-lives, so that after a time every generalization is "more history than science." This judgment underscores the validity of the naturalist's assumption that generalizations of the rationalistic variety are not possible because phenomena are intimately tied to the times and the contexts in which they are found.

TABLE 1
Scientific and Naturalistic Terms Appropriate to the Four Aspects of Trustworthiness

Aspect	Scientific Term	Naturalistic Term
Truth Value	Internal Validity	Credibility
Applicability	External Validity Generalizability	Transferability
Consistency	Reliability	Dependability
Neutrality	Objectivity	Confirmability

Yet these facts do not obviate the possibility that some transferability between two contexts may occur because of certain essential similarities between them. To determine the extent to which transferability is probable, one needs to know a great deal about both the transferring and receiving contexts, to have what Geertz (1973) has dubbed "thick description" about each.⁷ If the thick descriptions demonstrate an essential similarity between two contexts, then it is reasonable to suppose that tentative findings of Context A are also likely to hold in Context B (although, to be safe, an empirical test of that presumption should be made). For the naturalist, then, the concept analogous to generalizability (or external validity) is transferability, which is itself dependent upon the degree of similarity (fittingness) between two contexts. The naturalist does not attempt to form generalizations that will hold in all times and in all places, but to form working hypotheses that may be transferred from one context to another depending upon the degree of "fit" between the contexts.

Consistency. Within the rationalistic paradigm concern over consistency stems from the fact that instruments must produce stable results if those results are to be meaningful. Validity is a direct function of reliability; so, for example, it is easy to show that the validity of an instrument cannot exceed the square root of its reliability (Gulliksen, 1950). Reliability is thus not so much essential in its own right as it is a precondition for validity. The naturalist is also concerned with consistency, and for the same reasons; naturalistic instruments no more than rationalistic ones are likely to yield credible (the analog of valid) results if they do not exhibit consistency. But consistency is a trickier concept for the naturalist than for the rationalist. The latter, believing in a single reality upon which inquiry converges, can treat all instrumental shifts as error, but the

naturalist, believing in a multiple reality and using humans as instruments—instruments that change not only because of "error" (e.g., fatigue) but because of evolving insights and sensitivities—must entertain the possibility that some portion of observed instability is "real." Thus, for the naturalist, the concept of consistency implies not invariance (except by chance) but trackable variance—variance that can be ascribed to sources: so much for error, so much for reality shifts, so much for increased instrumental proficiency (better insights), and so on. The naturalist thus interprets consistency as dependability, a concept that embraces elements both of the stability implied by the rationalistic term *reliable* and of the trackability required by explainable changes in instrumentation.

Neutrality. Neutrality is commonly termed *objectivity* within the rationalistic paradigm. Objectivity is presumably guaranteed by methodology; If the methods are explicated, open to public scrutiny, replicable, and at least one step removed from direct investigator-subject contact, then objectivity is assured (that is, the biases of the investigator are effectively screened out). But of course methodology inevitably reflects the predispositions of the investigator. In physics, for example, it is no longer disputed that whether light is "proved" to be wave-like or corpuscular in nature depends entirely on which experiment one chooses to do: Young's double-slit experiment definitively shows light to be a wave, while Einstein's work with the photoelectric effect established its corpuscular nature without doubt (Zukav, 1979). In the social sciences, the cultural and ethnic biases that can be built into, for example, IQ measuring instruments, are well known. Naturalists are especially aware of this problem because they understand the multiple realities that one encounters (including multiple value systems) and the role that their own predispositions can play when they use themselves as instruments. Following the reasoning of Scriven (1972), naturalists shift the burden of neutrality from the investigator to the data, requiring evidence not of the certifiability of the in-

⁷The concept of "thick description" is not yet well explicated. For one effort to outline what such a description should contain, at least in an evaluation situation, see Guba and Lincoln (in press).

TABLE 2
The Rationalistic Treatment of Trustworthiness

<i>Inquiry can be affected by:</i>	<i>Which produce effects of:</i>	<i>To guard against which we:</i>	<i>In hope this action will lead to:</i>	<i>And produce findings that are:</i>
Masking or competing factors	Confounding	Control and/or randomize	Internal validity	Contamination-proof
Situational variations	Atypicality	Require probability sampling	External validity	Context-proof
Instrumental drift or decay	Instability	Replicate	Reliability	Inconsistency-proof
Investigator predilections	Bias	Insulate the investigator	Objectivity	Investigator-proof

investigator or his or her methods but of the confirmability of the data produced.

WHAT IS THE RATIONALISTIC MODE FOR DEALING WITH QUESTIONS OF TRUSTWORTHINESS?

The modes for dealing with the four areas of trustworthiness that have evolved within the rationalistic paradigm are summarized in Table 2. The four rows of the table correspond to the four areas: internal validity, external validity, reliability, and objectivity (see entries in the fourth column of the table). The columns of the table have headings that are intended to guide the reader through the cell entries. Begin by reading the heading of the first column, follow that by reading the entry in Cell 1, Row 1, read the heading of the second column and follow that by reading the entry in Cell 2, Row 1, and so on for the remaining columns and cells of Row 1. Repeat the process for each of the four rows. In effect, the table produces four sentences that recapitulate the conventional rationalistic response to the trustworthiness concerns. For example, the cells in Row 1 produce these sentences:

Inquiry can be affected by masking or competing factors,
which produce effects of confounding,

to guard against which we control and/or randomize,
in the hope that this action will lead to internal validity,
and produce findings that are contamination-proof.

Several observations may be made about Table 2:

1. The entries in the first column of the table—the detractors from “perfect” inquiry—are perceived as sources of error, elements extraneous to the thrust of the inquiry that represent intrusions or barriers to its orderly development. Methodology is constructed not so much to take account of these factors as to guard against the threats posed by their existence (Campbell & Stanley, 1963).

2. The aim of the counteractions noted in the third column of the table is to render the study proof against these threats. Findings must be certifiable as contamination-proof, context-proof, inconsistency-proof, and investigator-proof (Column 5). This awesome task is entrusted, within the rationalistic paradigm, to methodology—not to the expertise and insight of the investigator.

3. The appropriate methodologies that will produce such foolproof data are listed in the third column of the table. The actions noted there are the ideal or

"textbook" answers to the questions, "What have you done to guarantee internal validity? External validity? Reliability? Objectivity?" Investigators who can document that they have controlled or randomized all (noninformation) variables, selected and assigned subjects to treatments at random, replicated (established the reliability of) the instrumentation (even if only in the split-half sense), and insulated themselves from the phenomena by a layer of "objective" instrumentation have provided unassailable responses to those questions.

4. When a research proposal or report is examined for trustworthiness, or when the implementation of a design is monitored for procedures that will guarantee trustworthiness, it is the entries of Column 3 that are used as criteria. The examiner or monitor asks, "Did (or will) the investigators control or randomize? Did (or will)

they sample by probability methods? Did (or will) they replicate? Did (or will) they interpolate a layer of instrumentation between themselves and the phenomena?" If the answer (or putative answer) to those questions is "Yes," the proposal-process-report is deemed to be satisfactory insofar as trustworthiness is concerned.

WHAT IS THE NATURALISTIC MODE FOR DEALING WITH QUESTIONS OF TRUSTWORTHINESS?

The naturalistic mode of dealing with the four trustworthiness questions is summarized in Table 3, which follows the same format as Table 2. The reader will note that the middle column of Table 2 is represented in Table 3 by two columns, one detailing steps that can be taken during the inquiry, and the second, steps that can be taken after the inquiry is complete.

TABLE 3
The Naturalistic Treatment of Trustworthiness

<i>Inquiry can be affected by:</i>	<i>Which produce effects of:</i>	<i>To take account of which we:</i>		<i>In the hope these actions will lead to:</i>	<i>And produce findings that are:</i>
		<i>During:</i>	<i>After:</i>		
Factor patternings	Noninterpretability	Use prolonged engagement Use persistent observation Use peer debriefing Do triangulation Collect referential adequacy materials Do member checks	Establish structural corroboration (coherence) Establish referential adequacy Do member checks	Credibility	Plausible
Situational uniquenesses	Noncomparability	Collect thick descriptive data Do theoretical/purposive sampling	Develop thick description	Transferability	Context-relevant
Instrumental changes	Instability	Use overlap methods Use stepwise replication Leave audit trail	Do dependability audit (process)	Dependability	Stable
Investigator predilections	Bias	Do triangulation Practice reflexivity (audit trail)	Do confirmability audit (product)	Confirmability	Investigator-free

Following the logic of Table 2, Table 3 also can be interpreted to produce four sentences. For example, the cells in Row 1 produce these sentences:

Inquiry can be affected by factor patternings, which produce effects of noninterpretability, to take account of which we,
 while doing (*during*) the study, use prolonged engagement, persistent observation, and peer debriefing, do triangulation, collect referential adequacy materials, and do member checks,
 and *after* completing the study, establish structural corroboration or coherence, establish referential adequacy, and do member checks,
 in the hope that these actions will lead to credibility,
 and produce findings that are plausible.

The sentences in Table 3 are not self-interpreting; the following discussion will define and clarify their major terms and processes.

Credibility. Whereas rationalists are concerned with guarding against masking or competing factors (sources of error) that are said to confound the inquiry, naturalists wish to take account of the bewildering array of interlocking factor patterns that confront them and pose formidable problems of interpretation. The rationalists' solution to the problem is to abstract several variables of special interest, remanding the rest to the status of controlled or randomized variables. The naturalists' solution is to deal with the patterns in their entirety but to take certain actions that take account of the complexities.

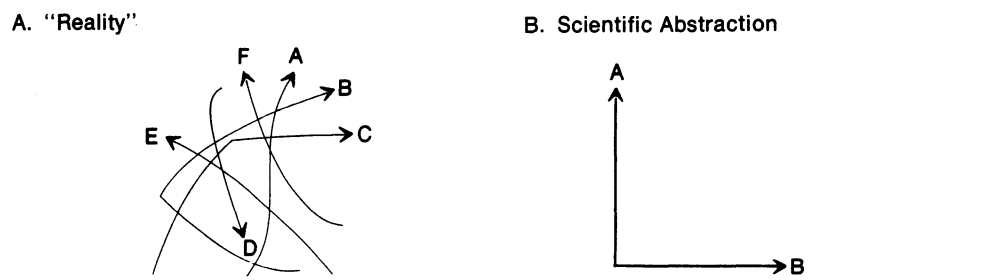
The situation is much as depicted in Figure 1A. The "reality" of the situation is that many factors, bearing a variety of rela-

tionships (correlations, rationalists would say) to one another, form a "whole" that cannot be understood if dismembered. Rationalists do just that, however: they single out several variables, and cast them into a design which by definition treats them as orthogonal (Figure 1B). To use Brunswik's (1955) terminology, rationalists *tie* certain variables in the design (that is, place them into a specific relationship determined not by nature but by the design) and *untie* certain others (that is, treat variables related in nature as though they were independent). This tying/untying is the ultimate effect of the rationalists' solution of controlling and/or randomizing.

Naturalists eschew this approach because they feel it does violence to the phenomena they seek to understand. Instead, they adopt certain other procedures which, while not as theoretically unassailable, nevertheless preserve the holistic situation. Methods that may be used during the study include:

- *Prolonged engagement at a site*, to overcome, so far as possible, distortions produced by the presence of researchers and to provide researchers the opportunity to test their own biases and perceptions, as well as those of their respondents. Spending an extended period at a site allows locals to adjust to the presence of researchers and to satisfy themselves that they do not constitute a threat. Researchers are given time to check their own developing perceptions; for example, by keeping journals. If, six months into the site engagement, the journals continue to refer to locals and to characterize their life styles in the same way, it is likely that the researchers are continuing to view the situa-

FIGURE 1
 "Reality" Versus Scientific Abstraction: Tying and Untying Variables



tion in terms of their early predispositions and have learned nothing from their presence at the site. Researchers must exercise caution, however, to avoid becoming over-involved with the respondents—what the anthropologists call “going native.”

- *Persistent observation*, in order to identify pervasive qualities as well as atypical characteristics. Extended interaction with a situation or a milieu leads inquirers to an understanding of what is essential or characteristic of it. At the same time they learn to eliminate aspects that are irrelevant while continuing to attend to those that, while atypical, are nevertheless critical (Eisner, 1979). Inquirers ought to be able to show that sufficient time was spent at the site to justify their characterization of it; their journals ought to reflect their wrestling with the question of what the pervasive qualities are.

- *Peer debriefing*, to provide inquirers the opportunity to test their growing insights and to expose themselves to searching questions. Inquirers ought regularly to detach themselves from the site and to seek out and interact with other professionals who are able and willing to perform the debriefing function; for example, faculty colleagues or members of a dissertation committee. Inquirers ought to expose their thinking to this “jury” of peers and to deal with whatever questions they may pose. Their journals and field activities ought to indicate timely redirection of the inquiry consistent with the critiques obtained during these debriefings.

- *Triangulation*, whereby a variety of data sources, different investigators, different perspectives (theories), and different methods are pitted against one another in order to cross-check data and interpretations (Denzin, 1978). For example, no item of information ought to be accepted that cannot be verified from at least two sources. When possible the research team should be divided so that the perceptions of several investigators can be compared. Different theories ought to be brought to bear on data to yield alternative explanations that can be tested. Different methods, for example, questionnaires, interviews, and documentary analyses,

should be used when possible. In the field, techniques of cross-examination should be used when reports from different informants are in conflict.

- *Collection of referential adequacy materials*, whereby documents, films, videotapes, audio recordings, and other “raw” or “slice-of-life” data items are collected against which findings and interpretations can later be tested (Eisner, 1979). So, for example, if the inquiry deals with the behavior of classroom teachers, videotapes of actual classrooms can be made and stored. Later, when it is asserted that teachers exhibit such and such behavior, that assertion can be tested by reference to the archives.

- *Member checks*, whereby data and interpretations are continuously tested as they are derived with members of the various audiences and groups from which data are solicited. The process of member checks is the single most important action inquirers can take, for it goes to the heart of the credibility criterion. Inquirers ought to be able to document both having made such checks as well as the ways in which the inquiry was altered (emerged or unfolded) as a result of member feedback.⁸

Methods that can be used after the study has been completed include:

- *Establishing structural corroboration or coherence*, that is, testing every datum and interpretation against all others to be certain that there are no internal conflicts or contradictions. Of course, particular data items may be in conflict because they come from different sources, represent different perspectives, and so on, but inquirers ought to be able to make an interpretation that explains these apparent contradictions. Interpretations should also take account of possible rival explanations and negative or deviant cases (Patton, 1980), a matter of special interest to critics who operate from the rationalistic point of view.

⁸An inquirer must be aware, of course, that audience members may choose to label the data and interpretations as noncredible for reasons of their own. A discussion of the ways in which audiences may dupe either themselves or an inquirer is beyond the scope of this paper. For a fuller discussion see Guba and Lincoln (in press) and Douglas (1976).

since the elimination of plausible rival alternatives constitutes their chief means for establishing internal validity. Finally, the overall report or case study should demonstrably exhibit coherence; that is, consistency, synchronism, logic, and being "all of a piece."

- *Establishing referential adequacy*, that is, testing analyses and interpretations made after completion of the field portion of the study against documents, recordings, films, and the like that were collected or especially produced for this purpose while the study was under way. Researchers should, within the limits of time and fiscal constraints, have collected referential adequacy materials during the study. These materials can be used to conduct empirical tests; for example, by constituting panels or juries of "experts" to test conclusions against these raw materials.

- *Member checks*, that is, testing the overall report or case study with source groups before casting it into final form. These checks are like those already described, but are carried out after completion of the study rather than during it.

Transferability. Naturalists eschew generalizations on the grounds that virtually all social/behavioral phenomena are context-bound. It is not possible, they believe, to develop "truth" statements that have general applicability; rather, one must be content with statements descriptive or interpretative of a given context—idiographic or context-relevant statements. During a study naturalists will:

- *Do theoretical/purposive sampling*, that is, sampling that is *not* intended to be representative or typical (such a purpose focuses the investigator on similarities and makes sense only when one is trying to generalize) but that *is* intended to maximize the range of information uncovered. The nature of the sampling process is governed by emergent insights about what is important and relevant. Naturalistic investigators ought to be able to demonstrate how the samples they selected met this criterion—for example, that successive interview subjects were selected by asking

each respondent to nominate someone whose point of view is as different as possible from his or her own.⁹

- *Collect "thick" descriptive data* that will permit comparison of this context to other possible contexts to which transfer might be contemplated (Geertz, 1973). If transferability depends upon a match of characteristics it is incumbent on investigators to provide the information necessary to test the degree of fittingness.

After the study is completed naturalists will:

- *Develop thick description* of the context in order to make judgments about fittingness with other contexts possible. Inquirers ought to make available, as an appendix to their reports or in a supplement available to interested parties, a full description of all contextual factors impinging on the inquiry.

Dependability. Naturalists are concerned with the stability of data, but must make allowance for apparent instabilities arising either because different realities are being tapped or because of instrumental shifts stemming from developing insights on the part of the investigator-as-instrument. Two steps that can be taken during a study that parallel the replication steps typically advocated by the rationalists are:

- *Overlap methods*, one kind of triangulation process, whereby different methods are used in tandem (Campbell & Stanley, 1963; Webb, Campbell, Schwartz, & Sechrest, 1966). This approach is usually advocated to overcome invalidities in individual methods; two or more methods are teamed in such a way that the weakness of one is compensated by the strengths of another. But it is clear that if similar results are found using different methods the case for stability is also strengthened (Guba, 1978; Guba & Lincoln, in press). This so-called "multiple-operations" inquiry simultaneously undergirds the case for credibility and stability. Inquirers ought to be able to report on the use of multiple methods and demonstrate that these

⁹A useful discussion of theoretical sampling may be found in Glaser and Strauss (1967).

methods were selected because they were complementary (among other possible reasons).

- *Stepwise replication*, analogous to the "split-half" reliability of tests, in which two separate research teams (the original team split into halves) deal separately with data sources that have also been divided into halves. Because of the unfolding nature of naturalistic designs, however, the two teams cannot be permitted to pursue the inquiry to its end before results are compared. Provision must be made for communication between the teams at important milestone points, perhaps even on a daily basis, in order to cross-check developing insights and to decide on appropriate next steps. These communication sessions should be adequately documented.

A quite different approach to the more conventional steps outlined above is built on the metaphor of the fiscal auditor (Guba, 1978; Guba & Lincoln, in press). When, for example, an auditor from Price, Waterhouse is called in to audit the books of the General Electric Company, he or she takes on two responsibilities: (1) to examine the method of accounting by which the GE books are kept, to verify that the method used is among those generally accepted by the accounting profession (that is, to make sure that no "creative accounting" has taken place), and (2) to certify that the "bottom line" is correct; that is, that supportive documents (data) exist to support every entry and the addition (interpretation) is correct. In relation to dependability, it is the first of these applications that is relevant; that is, the examination of method. Thus, naturalists will, during a study:

- *Establish an "audit trail"* that will make it possible for an external auditor to examine the processes whereby data were collected and analyzed, and interpretations were made. The audit trail takes the form of documentation (the actual interview notes taken, for example) and a running account of the process (as in the form of an investigator's daily journal).

After completion of a study naturalists will:

- *Arrange for a "dependability" audit* to be done by an external auditor—someone competent to examine the audit trail and to comment on the degree to which procedures used fall within generally accepted practice. Such a dependability audit, it should be noted, deals primarily with the *processes* of the inquiry.

Confirmability. We have noted the shift by naturalists away from the concept of investigator objectivity toward the concept of data (and interpretational) confirmability. In the interest of confirmability, two of the steps naturalists can take during a study are:

- *Triangulation*, as already noted in relation to credibility—that is, collecting data from a variety of perspectives, using a variety of methods, and drawing upon a variety of sources so that an inquirer's predilections are tested as strenuously as possible. So far as it is feasible to do so, other investigators should also be employed. Research teams can be constituted so as to balance out predispositions; for example, by seeing to it that both rationalistic and naturalistic perspectives are represented (Guba & Lincoln, in press). An inquirer should provide documentation for every claim from at least two sources; alternative possibilities and negative instances should be ruled out, and so on.

- *Practicing reflexivity*, that is, to "intentionally reveal to his [or her] audience the underlying epistemological assumptions which cause him [or her] to formulate a set of questions in a particular way, and finally to present his [or her] findings in a particular way" (Ruby, 1980). One indispensable technique in support of practicing reflexivity is to keep a continuing journal in which introspections are recorded on a daily basis (Spradley, 1979); these introspections can also be tested during the peer debriefings already mentioned. Reinharz (1979) calls attention to the fact that reports of research typically include discussion only of the problem and the method; she suggests that it is equally important to discuss the inquirer and to document shifts and changes in his or her orientation.

After completion of the study naturalists will:

- *Arrange for a confirmability audit* that undertakes the second of the two auditing tasks described above—that is, an audit certifying that data exist in support of every interpretation and that the interpretations have been made in ways consistent with the available data. This type of audit is concerned primarily with the products of the inquiry, and requires extensive documentation. The audit can be performed at the same time by the same external agent commissioned to perform the dependability audit, as is usually the case in the parallel fiscal audit.

Several observations may be made about Table 3:

1. The entries in the first column of the table, unlike the case of Table 2 entries, are not viewed as "error" but as the "natural state of things," as reflections of reality in the minds and lives of respondents. They are therefore not to be guarded against but to be taken account of. Naturalists' strategies for establishing trustworthiness are aimed in that direction.

2. The entries in the middle two columns contains a naturalist's "textbook answers" to questions of trustworthiness just as did Column 3, Table 2 entries for the rationalistic paradigm. If a naturalist is challenged on the issue of credibility, for example, a proper defense is to respond, "I used prolonged engagement, I used persistent observation . . . I did member checks." But we must note one major difference between Table 2 and Table 3. Table 2 entries are not only "textbook answers" but in fact constitute an unassailable defense. Given the assumptions of the rationalistic paradigm, the steps outlined in Column 3 of Table 2 are theoretically complete steps that guarantee the study against threats to trustworthiness. The same claim cannot be made for Table 3—about all one can say is that to take these steps increases the probability of the study's trustworthiness.¹⁰ In that sense the

naturalistic theory of trustworthiness is an incomplete one—one cannot muster evidence that will compel another to accept the trustworthiness of the study but only evidence that will persuade the other of its relative trustworthiness. But this situation neither surprises nor dismays naturalists—that kind of indeterminacy is what they expect of the "real" world. Their response to someone who cannot tolerate that degree of ambiguity is simply to say, "Whoever promised you a rose garden?"

3. The responses outlined in the middle columns of Table 3 are not all equally weighty in persuading a critic of the trustworthiness of naturalistic inquiry. Some are *sine qua nons*, some merely *desiderata*. For example, it is inconceivable that one would be persuaded of the trustworthiness of a study that involved no triangulation and no member checks. Prolonged engagement, persistent observation, peer debriefing, and collecting referential adequacy materials might appear less necessary (however desirable). It is likely that triangulation and member checks (for credibility), thick description (for transferability), leaving an audit trail (for dependability), and triangulation and practicing reflexivity (for confirmability) are the minimums that should be required of naturalistic investigators.

4. Despite the indeterminacy that continues to surround the naturalistic study when questions of trustworthiness are raised, it is still the case that the entries in the middle columns of Table 3 constitute, at this point in time and thought, the best available formulation of criteria for judging naturalistic inquiries. Thus, when a naturalistic study is to be judged, it is these criteria that ought to be brought to bear.

IMPLICATIONS

If the preceding analysis has meaning, a number of implications can be drawn from it:

1. *Naturalistic inquiry has its own set of criteria for adequacy.* It is inappropriate to apply the rationalistic criteria of Table 2 under any circumstances. To suggest, for example, that a naturalistic study is unac-

¹⁰There is an interesting parallel between the determinacy of classical physics and the rationalistic paradigm on the one hand, and the indeterminacy and probabilistic nature of quantum physics and the naturalistic paradigm on the other hand.

ceptable because controls were not instituted, subjects were not randomly selected, instrumental results were not replicated, or the investigator was not properly objective is simply unjustified. But it is also inappropriate to apply, without thought, other, broader criteria; for example, those proposed by Schwen (1977) in an earlier ERIC/AVCR (now ECTJ) Annual Review Paper. While some of Schwen's criteria might be acceptable to naturalists, others are not; each must be considered independently.¹¹

2. *The proposed criteria, like scientific criteria, have utility at several stages in the inquiry process:*

- For making a priori judgments as in the case of proposals or designs (insofar as that term is appropriate). The proposal or design should indicate what an inquirer proposes to do to satisfy each of the criteria suggested in Table 3.
- For monitoring inquiry procedures. Investigators can utilize the criteria of Table 3 to guide their field activities and to impose checks on themselves to be certain that they are meeting criterial requirements.
- For making ex post facto judgments as in the case of reports or case studies. The report or case study should include statements about what an inquirer actually did to satisfy each of the criteria suggested in Table 3. An important part of such reports or case studies must be the results of dependability and confirmability audits.

¹¹Schwen did not limit himself to the trustworthiness criteria but took up the question of how to judge inquiry as professional scholarship. In the context of the present paper, some of Schwen's criteria would be readily accepted by naturalists: Inquiry should be publicly verifiable; it should be directed toward an extension of knowledge; it should demonstrate creative exploration; and it should make intuitive good sense. But others of Schwen's criteria would be less acceptable: Inquiry should be disciplined insofar as that term implies a preordinate design; it should be directed toward generalizability; it should be incorporated within a conceptual structure or theoretical framework (unless that structure of framework is grounded in the inquiry itself); and the instrumental problem should be a representation of some larger unresolved logical problem (which responds to the rationalist's need for generalizability).

3. *The statement of criteria is not equivalent to the statement of decision rules for applying these criteria.* It is one thing to suggest that triangulation is needed, for example, and quite something else to say how much, or what type, of triangulation will suffice to establish a minimal level of trustworthiness. It is one thing to specify that a dependability audit be done and quite another to establish the precise processes that constitute an adequate audit. It seems likely that the development of decision rules will be an empirical matter; only through efforts to apply the criteria of Table 3 will the field come to some understanding of what decision rules make sense. What we have here is a situation parallel to asking, "Is a reliability of .65 sufficient to establish the adequacy of a paper-and-pencil test?" or, "Is a questionnaire return rate of 46 percent adequate?" These questions can be answered only through experience.

4. *The use of the naturalistic paradigm is fraught with special risks for an investigator.* Because of the open-ended, initially design-less, emergent quality of naturalistic inquiry, it cannot be warranted in advance to the same extent as can rationalistic inquiry. Rationalists can propose to do random sampling, for example, and can detail just how they will go about doing it. A sponsor or client, a departmental chairman or college research committee, or a dissertation committee will have no difficulty in deciding whether the proposed procedure is rigorously correct and whether it will or will not produce the desired outcomes. When naturalists, by contrast, propose to do theoretical or purposive sampling, the issue is very much in doubt. Persons who must make judgments about the likelihood that the proposed procedure will work necessarily feel less certain about their judgments; they may feel that they are being asked to extend *carte blanche* to the researcher.

This ambiguity poses many problems. A funding source having to choose between a naturalistic inquiry and a rationalistic one is more likely to go with the latter because the outcomes seem more certain. Dissertation committees, already uncertain about the legitimacy of naturalistic

inquiry, are likely to reject a naturalistic dissertation because they are unwilling to risk their status as knowledgeable critics of research. In all instances the naturalistic approach is likely to be tarred with the brush of "sloppy research"—research that cannot be better specified than *that* must perforce be inadequate. There are thus a good many barriers and constraints to inhibit an investigator from undertaking a naturalistic study in the first place.

There is little that a naturalistic inquirer can do about such attitudes at the moment other than to accept them as part of the *res gestae*. It is the author's hope that the specification of criteria in this paper, however inadequate at this primitive stage, will serve to stimulate discussion about these problems and to increase the tolerance of the professional community for those of their colleagues who elect to conduct their studies by these newer rules.

5. *The proposals of Table 3 should not be reconstituted into an orthodoxy.* One of the major difficulties in proposing a new paradigm is that the old is so entrenched—it is no longer a way to do inquiry but *the* way. Kaplan (1964) has called the conceptualizations inquirers produce about the ways they do inquiry *reconstructed logic*. At best reconstructed logics are afterthoughts that describe what the inquirer believes he did; most often they do not adequately describe what the inquirer actually did (what Kaplan terms the logic-in-use). Now reconstructed logics have many uses: to train the novice, to facilitate communication among practitioners; to provide checkpoints against which inquirers can test themselves. But—and most emphatically—they are not prescriptions of how inquiry must be done. When reconstructed logics are allowed to become orthodoxies, inquirers are reduced to becoming true believers, a posture hardly consonant with the open position they typically espouse. The level of this paper, it should be recalled, is primitive; I hope new and better criteria will quickly replace those suggested here. It is dubious whether the "perfect criteria" will ever emerge; until then, humility in asserting that a "new and truer path to knowledge" has been found will be wise.

REFERENCES

- Brunswik, E. Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 1955, 62, 193–217.
- Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago, Ill.: Rand McNally, 1963.
- Cronbach, L. J. Beyond the two disciplines of scientific psychology. *American Psychologist*, 1975, 30, 116–27.
- Denzin, N. K. The logic of naturalistic inquiry. In Norman K. Denzin (Ed.), *Sociological methods: A sourcebook*. New York: McGraw-Hill, 1978.
- Douglas, J. D. *Investigative social research*. Beverly Hills, Calif.: Sage Publications, 1976.
- Eisner, E. W. *The educational imagination*. New York: Basic Books, 1979.
- Geertz, C. *The interpretation of culture*. New York: Basic Books, 1973.
- Glaser, B. G., & Strauss, A. L. *The discovery of grounded theory*. Chicago, Ill.: Aldine Publishing, 1967.
- Guba, E. G. Toward a methodology of naturalistic inquiry in educational evaluation. *CSE monograph series in evaluation*, No. 8. Los Angeles, Calif.: Center for the Study of Evaluation, UCLA, 1978.
- Guba, E. G. Naturalistic inquiry. *Improving Human Performance Quarterly*, 1979, 8, 268–76.
- Guba, E. G., & Lincoln, Y. S. *Effective evaluation*. San Francisco, Calif.: Jossey-Bass, in press.
- Gulliksen, H. *Theory of mental tests*. New York: Wiley, 1950.
- Kaplan, A. *The conduct of inquiry*. San Francisco, Calif.: Chandler Publishing, 1964.
- Patton, M. Q. *Qualitative evaluation methods*. Beverly Hills, Calif.: Sage Publications, 1980.
- Polanyi, M. *Personal knowledge*. New York: Harper & Row, 1958.
- Reinharz, S. *On becoming a social scientist*. San Francisco, Calif.: Jossey-Bass, 1979.
- Ruby, J. Exposing yourself: Reflexivity, anthropology, and film. *Semiotica*, 1980, 30, 153–79.
- Schwen, T. M. Professional scholarship in educational technology: Criteria for judging inquiry. *AV Communication Review*, 1977, 25, 5–24.
- Scriven, M. Objectivity and subjectivity in educational research. In Lawrence G. Thomas (Ed.), *Philosophical redirection of educational research*, 71st Yearbook of the National Soci-

- ety for the Study of Education, Part 1. Chicago, Ill.: University of Chicago Press, 1972.
- Scriven, M. Maximizing the power of causal investigations: The *modus operandi* method. In G. V. Glass (Ed.), *Evaluation studies review annual*, vol. 1. Beverly Hills, Calif.: Sage Publications, 1976.
- Spradley, J. P. *The ethnographic interview*. New York: Holt, Rinehart, & Winston, 1979.
- Stake, R. E. *Evaluating the arts in education*. Columbus, Ohio: Charles E. Merrill, 1975.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. *Unobtrusive measures*. Chicago, Ill.: Rand McNally, 1966.
- Wolf, R. L. *An overview of conceptual and methodological issues in naturalistic evaluation*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, Calif.: April 1979.
- Wolf, R. L., & Tymitz, B. Toward more natural inquiry in education. *CEDR Quarterly*, 1977, 10, 7-9.
- Zukav, G. *The dancing Wu Li masters: An overview of the new physics*. New York: Bantam, 1979.

CALL FOR CONVENTION PAPERS

This Research and Theory Division (AECT) announcement is an official call for papers for presentation at the 1982 AECT Convention, to be held in Dallas.

Types of Papers

The following types of papers are being sought:

- Research Papers
- Conceptual Papers
- Symposia Papers
- Papers Describing Research in Progress

Presentations

Presentations will be 15 minutes in length, with three to five minutes allowed for questions. Discussants previously identified and familiar with the paper will be present to raise questions and to stimulate dialogue and discussion after the presenta-

tion. (Authors of accepted papers must provide discussants with two copies of their completed papers by April 1, 1982.)

Proposals

Proposals for papers must be submitted by October 30, 1981. Proposals should be approximately 1,000 words long and should summarize the major elements of the paper.

How to Apply

A special cover sheet must accompany the proposal. For more information and a copy of the cover sheet, write to:

James Sucy, Manager
Educational Market Services
Eastman Kodak Company
343 State Street
Rochester, NY 14650