Disciplines of Inquiry
in Education: An Overview

LEE S. SHULMAN
Michigan State University

Few works in the English language are as rich as Shakespeare's *Hamlet*. One phrase is of particular interest. Hamlet is in deep grief and despair over the recent death of his father, King of Denmark. In his melancholy he has been acting rather strangely and many have called him mad. Yet, Polonius observes of Hamlet, "Though this be madness, yet there is method in it."

What does Shakespeare wish to convey with this phrase? How can the apparent lack of coherence or sanity of Hamlet's behavior be characterized by method? To assert that something has method is to claim that there is an order, a regularity, obscure though it may be, which underlies an apparent disorder, thus rendering it meaningful. Method is the attribute which distinguishes *research* activity from mere observation and speculation.

When adversaries argue about the nature of the world or the best approach to some particular human endeavor, we typically find ourselves evaluating their respective claims through examining the methods they used to reach their conclusions. There are few subjects that generate as much passion among scientists as arguments over method. This is not surprising, since scholars who agree on matters of method can

pursue research questions in a parallel fashion and then argue over the results of their respective investigations. However, if they do not agree even on some matters of research method, then their findings are likely to be incommensurable. There will be no way to properly compare one inquiry with the other. It is for this reason that major controversies in educational research so frequently focus on problems of research method. What is the role of research methodology in educational research? How can we tell proper from improper uses of research methods? To answer these questions, we must turn to a central concept in educational research methodology—disciplined inquiry.

Method and Disciplined Inquiry

Educational researchers are typically eager to distinguish their work from other forms of discourse which, for them, cannot lay claim to being research. Take for example the following statement from the preface of Cremin's (1961) prize-winning history of American progressive education, *The Transformation of the School.*

There is currently afoot a simple story of the rise of progressive

education, one that has fed mercilessly on the fears of anxious parents and the hostilities of suspicious conservatives. In it John Dewey, somewhat in the fashion of Abou Ben Adhem, awakes one night with a new vision of the American school: the vision is progressive education. Over the years, with the help of a dedicated group of crafty professional lieutenants at Teachers College, Columbia University, he is able to foist the vision on an unsuspecting American people. The story usually ends with a plea for the exorcising of this devil from our midst and a return to the ways of the fathers. This kind of morality play has always been an influential brand of American political rhetoric, used by reformers and conservatives alike. But it should never be confused with history! (p. vii)

Cremin forcefully draws the distinction between doing history and engaging in political rhetoric. Clearly, he claims, the results of the two forms of discourse must be treated with different degrees of respect and credibility. "Real history" should be given far greater credence than mere political rhetoric. How is one to distinguish between the two? I would suggest that, while not entirely a matter of method, the historian would distinguish his work from that of the rhetorician by the ways observations are collected, evidence is marshalled, arguments are drawn, and opportunities are afforded for replication, verification, and refutation.

When we speak of research, we speak of a family of methods which share the characteristics of *disciplined inquiry.* Cronbach and Suppes (1969) attempted to define disciplined inquiry a number of

years ago in a monograph prepared with the collaboration of their colleagues in the National Academy of Education. Here are some of the definitions of disciplined inquiry they suggest:

> Disciplined inquiry has a quality that distinguishes it from other sources of opinion and belief. The disciplined inquiry is conducted and reported in such a way that the argument can be painstakingly examined. The report does not depend for its appeal on the eloquence of the writer or on any surface plausibility. (p. 15)

Hence, perhaps, a major difference between disciplined inquiry and political rhetoric is that disciplined inquiry does not depend on eloquence or surface plausibility alone. Cronbach and Suppes continue:

> Whatever the character of a study, if it is disciplined the investigator has anticipated the traditional questions that are pertinent. He institutes control at each step of information collection and reasoning to avoid the sources of error to which these questions refer. If the errors cannot be eliminated he takes them into account by discussing the margin for error in his conclusions. Thus, the report of a disciplined inquiry has a texture that displays the raw materials entering the argument and the logical processes by which they were compressed and rearranged to make the conclusion credible. (pp. 15-16)

That definition of disciplined inquiry could be misconstrued to imply that the appropriate application of research methods in education always leads to a sterile, ritualized and narrowly conceived form of investigation. This is not the case. As Cronbach and Suppes observe subsequently,

> Disciplined inquiry does not necessarily follow well established, formal procedures. Some of the most excellent inquiry is free-ranging and speculative in its initial stages, trying what might seem to be bizarre combinations of ideas and procedures, or restlessly casting about for ideas. (p. 16)

What is important about disciplined inquiry is that its data,

arguments and reasoning be capable of withstanding careful scrutiny by another member of the scientific community.

If it is clear what constitutes disciplined inquiry and there is little disagreement regarding the need for research methods to be consistent with the standards of disciplined inquiries, why should this field be so filled with controversy? There are several reasons.

First, scientific inquiries cannot involve mere recitation of the "facts of the case." Indeed, inquiry demands the selection of a particular set of observations or facts from among the nearly infinite universe of conceivable observations. Just as in a court of law the legal adversaries may disagree profoundly about the relevance of a piece of evidence or the warrant to be given to the conclusions drawn from each other's reasoning, so in disciplined inquiry in education there is often lack of consensus about the grounds, the starting points, for chains of reasoning.

There is another, even more serious source of disagreements about method. Disciplined inquiry not only refers to the ordered, regular, or principled nature of investigation, it also refers to the disciplines themselves which serve as the sources for the principles of regularity or canons of evidence employed by the investigator. What distinguishes disciplines from one another is the manner in which they formulate their questions, how they define the content of their domains and organize that content conceptually, and the principles of discovery and verification that constitute the ground rules for creating and testing knowledge in their fields. These principles are different in the different disciplines.

A major reason why research methodology in education is such an exciting area is that education is not itself a discipline. Indeed, *education is a field of study,* a locus containing phenomena, events, institutions, problems, persons, and processes, which themselves constitute the raw material for inquiries of many kinds. The perspectives and procedures of many disciplines can be brought to bear

on the questions arising from and inherent in education as a field of study. As each of these disciplinary perspectives is brought to bear on the field of education, it brings with it its own set of concepts, methods, and procedures, often modifying them to fit the phenomena or problems of education. Such modifications, however, can rarely violate the principles defining those disciplines from which the methods were drawn.

Applications of Research Methods: Some Examples

It is important to recognize that differences in method are not merely alternative ways of reaching the same end or answering the same questions. What distinguishes methods from one another, usually by virtue of their contrasting disciplinary roots, is not only the procedures they employ, but the very types of questions they tend to raise. This point might be best understood if I take an area of educational inquiry and describe how questions would be asked and studies conducted from the perspectives of different forms of disciplined inquiry in that field of study. Each of the examples I draw will be credible pieces of research; that is, forms of disciplined inquiry. This exercise will illustrate the variety of forms of research method that can be employed in a disciplined manner in the same domain of inquiry.

One of the most important areas of educational research is the study of reading. In these days of "back to basics" in education, there is no basic more basic than reading. Millions of dollars and the time of thousands of individual investigators are invested in research to help us understand more about the teaching and learning of reading. Research is conducted in order to become smarter about certain matters, in this case, the teaching and learning of reading. What do we wish to know? What kinds of questions ought we to ask about language, reading, and learning? What kinds of reading research are possible and what can we learn from each?

One reasonable question is, "What makes some people suc-

cessful readers and others unsuccessful?" How can you predict which sorts of people are going to have difficulty learning to read, in order, perhaps to institute preventive measures before serious damage has been done? In this sort of research, one would collect a variety of measures on individuals, including measures of their performance on a number of tasks, their demographic or personal characteristics, aspects of their backgrounds and anything else that could conceivably assist in accurate prediction of the likelihood of reading difficulty or failure. An investigator would then employ the techniques of correlation and regression to investigate the relationships between those predictors and sets of useful outcome measures of reading performance for students of various ages. Correlation as a statistical procedure used to determine whether two variables are related, or how much they are related. The approach would be quantitative and would involve no intervention or manipulation other than that required to administer the instruments needed to collect the necessary data. In general, correlational research attempts to describe the relationships among naturally occurring variables or phenomena, *without* attempting to change them.

Another investigator might now say, "I'm not really interested in predicting reading failure. I want to identify the best possible methods for teaching reading to all youngsters, irrespective of their backgrounds or aptitudes." Such an individual is unlikely to be satisfied with research methods that correlate attributes of individuals with concurrent or subsequent reading performance. This individual will be inclined to design experimental studies. Individuals or groups are systematically assigned to contrasting methods of reading instruction. The effects of these contrasting methods are then compared by testing the reading performance of those who have been taught. This approach involves experimental methods which contrast strikingly with those of correlational research. Naturally, there are times when the degree of control over the assignment of individuals or groups to treatments is not as great as may be theoretically desirable. We may, for example, wish to contrast two schools which are using very different reading programs. Since pupils were not originally assigned to those schools at random, this cannot be considered a "true" experiment. In such cases we see researchers use other methods which attempt to identify which treatment was best without the benefits of random assignment. These are often called "quasi-experimental" procedures. Like "near beer," they are not the real thing, but they can come close.

Yet another investigator may say that neither predicting reading failure nor identifying the best methods of teaching reading constitute the questions of interest for him. Instead his questions may be, "What is the general level of reading performance across different age, sex, social or ethnic groups in the population?" "Where do the most significant areas of reading success and failure occur?" "What are the reading habits of particular groups in the general population?" This investigation will best be conducted through a variety of survey techniques measuring reading performance or questioning reading practices. The work of the National Assessment of Educational Progress or of the International Education Association studies of cross-national achievement exemplify this approach. Once again, different procedures are employed to ask different questions and to solve different problems.

In the cases I have described thus far, the significant questions concern how well or how much reading ability has been gained or developed. Thus there are comparisons between alternative methods of teaching reading or among different individuals or cohorts of students learning to read. Quite another sort of question can be asked about reading. There are many times when we wish to know not how many or how well, but simply how. How is reading instruction carried on? What are the experiences and perceptions of teachers and students as they engage in the teaching and learning of reading? What is the underlying or explicit system of rules by which this complex activity is accomplished?

Although at first blush this might seem a much less powerful form of question than the quantitative questions which preceded it, this is not necessarily the case. Some of the most important and influential investigations in the history of social science have been of that form. Perhaps an example can illustrate that point.

When Binet and Simon were asked to develop a better method for identifying the children in the public schools of Paris who could profit from special education programs, they responded by creating the individual intelligence test. The goal of their research and development was to improve the precision with which one could measure the differences in intellectual ability among persons.

Nearly 20 years later, a young Swiss associate of Simon, Jean Piaget, became intrigued with a very different sort of question about human intelligence. He asked, "What does intelligence look like and how does it develop?" He was most concerned with the common elements characterizing the intellectual performance of all individuals at a given stage of development, rather than the levels of performance that distinguished among them. He was attempting to answer questions about shared regularities rather than measuring systematic differences.

Similarly, an individual interested in investigating the game of golf may decide to focus on differences in performance among golfers. What distinguishes good golfers from poor golfers? The study can be conducted experimentally, through contrasting alternative methods of training golfers. It can be accomplished correlationally by examining the attributes of poorly and well-scoring golfers through use of everything from videotape analyses to measures of age, experience, and social characteristics. But a very different question would be, "How does one play the game of golf?"

What are the functional rules of the game? In this case the investigator is interested in understanding the common elements or regularities shared by all golfers, whether they are national champions or weekend duffers.

To continue these examples about reading, there are investigators attempting to understand how reading instruction is accomplished in the classroom in general. They tend to use the methods of case study as they document or portray the everyday experiences of teachers and students in the teaching and learning of reading. In much case-study work, there is a general assumption that American public schools are very similar to one another as institutions. Therefore, individual experiences of learning to read will not be enormously different from one setting to another. In other case-study work, the assumption may be that "average" reading development does not exist. These researchers wish to document the dramatic diversity among individuals in the rate, sequence, and character of their development of reading competence.

These studies are likely to focus on only one classroom or school, a small number of them at most. Depending on the orientation of the researcher, the portrayals could emphasize the social character of learning to read in a classroom group, the intellectual and emotional experiences of individual children struggling to master the intricacies of reading, or even the manner in which individual children acquire the implicit rules for turn taking and status attainment in the classroom. Data gathering can include detailed prose descriptions written longhand on yellow pads, videotaping classroom episodes and analyzing their contents exhaustively, interviewing teachers and students to discover their reactions, perceptions, or expectations in classrooms and collecting examples of work produced by teachers and students for careful review and interpretation.

The disciplines from which these methods draw their rules of discovery and verification are typically anthropology, ethology,

linguistics, or particular subfields of sociology, such as symbolic interaction. These contrast sharply with the disciplinary roots of the more quantitative approaches, predominantly psychology, agriculture, genetics, and quantitative sociology, such as demography.

A philosopher approaching the problem of research in reading might raise yet another set of questions. He or she might examine the kinds of inquiry just described and observe that the *concept* of reading has not been adequately defined. What does it mean to be able to read? Do we denote by the term "reading" the ability to recognize the correspondence between visible symbols and sounds in isolation, mere word identification? Do we imply the ability to comprehend written prose, and, if so, at what level of sophistication or subtlety? For example, does someone who knows how to read have the ability to detect the difference between assertion and irony in a prose passage? Analysis of the meaning of the reading process affects the kinds of tests and measurements of reading achievement that are constructed. What we choose to define as reading will be important whether pursuing predictive studies of reading failure, experimental studies of reading instructional methods, or general surveys of reading performance. A philosopher would conduct inquiries into the nature of the reading process that would entail quite different research procedures from those of the other investigators. These analytic procedures would be disciplined by the rules of evidence proper to philosophy.

Similarly, questions of what distinguishes readers from nonreaders can be approached historically. As soon as someone attempts to answer the question "What proportion of the U.S. population is illiterate?" the ambiguity of the definition of "literacy" becomes apparent. How well must a person read and write to be considered literate? How has that definition changed for societies with contrasting economic systems, religious orientations, sex-role

prescriptions, or social-class hierarchies? A careful historical analysis can help account for both the conditions that foster increased literacy among members of a society, and the possible consequences of illiteracy for those members.

I have attempted in the examples just presented to illustrate the variety of ways in which types of research method can be applied to a topic of inquiry in education—reading instruction. Moreover, I have tried to indicate that the alternative methods not only approach the doing of research differently, but, by and large, ask different questions, and, hence, generate quite different answers. This is hardly surprising and surely not disturbing. The need for a multiplicity of methods was recognized centuries ago, perhaps most eloquently by Aristotle, who, in the introduction to his treatise *De Anima (On the Soul),* observes,

> It might be supposed that there was some single method of inquiry applicable to all objects whose essential nature we are endeavoring to ascertain . . . in that case what we should seek for would be this unique method. But if there is no single and general method for solving the question of essence, our task becomes still more difficult; in the case of each different subject we shall have to determine the appropriate process of investigation. (*Ethics,* I:1) (McKeon pp. 145-146)

Generalizability of Research

However different the objects of investigation and the goals of inquiry, there are certain problems shared by all research methods. These problems include the *generalizability* of findings; that is, the degree to which findings derived from one context, or under one set of conditions may be assumed to apply in other settings or under other conditions. Although there may be disclaimers from some research practitioners, all researchers strive for some degree of generalizability for their results. They are rarely content to have the research they have conducted generate under-

standing that is relevant only to the particular cases that were observed. There are several forms of generalization. The most frequently discussed is generalization from the particular sample of individuals who are tested, taught or observed in a given study to some larger population of individuals or groups of which they are said to be representative. For example, if we conduct a study of reading comprehension with third graders in Philadelphia, can we generalize our results to third graders all over the country? Or must we limit our generalizations to children of certain social and economic backgrounds, ability levels, and the like?

The second form of generalization is from the particular tasks or settings in which a piece of work is conducted to that population of tasks or settings that the research situation is claimed to represent. For example, we may contrast phonics and whole-word approaches to reading instruction using two particular sets of books or methods. If we find one approach consistently superior, can we generalize these findings to *all* phonics and whole-word methods? Or must we limit our generalizations to those particular teaching materials alone?

While both types of generalizability are important, much more has been written about the first kind, generalizability across people, than about the second, generalizability across situations. We shall see that the two have certain elements in common. In classical statistics, the argument was made that if one samples randomly from a population in making certain measurements or conducting certain experiments, inferences can then properly be drawn to the entire population from which the random sample was taken. Unfortunately, it is rarely the case that investigators truly sample randomly from a total population to which they might ultimately wish to generalize. A truly random sample is one in which each individual in the population has an equal chance of appearing. In a now classic paper, Cornfield and Tukey (1956) have argued that this is

never the case. Indeed, we sample as best we can and then make a case for the subsequent claims of generalizability. To use their metaphor, we must then build an inferential bridge between the particular groups of people whom we studied directly in our research and those other groups concerning whom we wish to generalize. We do so by documenting as comprehensively as necessary the characteristics of the individuals whom we have studied and the procedures we have used. Then, the reader can examine our documentation and critically evaluate whether our claims of generalizability are warranted. More specifically, the reader must judge whether the findings we report for the individuals whom we have studied should be considered applicable to any other group of individuals regarding whom our reader might be interested.

Cornfield and Tukey's concept of bridge building extends fruitfully to other aspects of generalization as well. When we report on a setting or a task we must be equally careful to document in detail its characteristics so that readers who are as concerned about the generalizability of our task characteristics as they are the generalizability of our sample can make the appropriate inferences.

Finally, we can now see that those who perform case studies are confronted with a problem of generalizability that is not different in kind from that confronted by their quantitative colleagues. To claim that one is conducting a case study requires that an answer be provided to the question, "What is this a case of?" Not every description is a case study. It may be a description of a singular individual or event. To claim that something is a case study is to assert that it is a member of a family of individuals or events of which it is in some sense representative. In much the same way that the reader of a quantitative study must build his Cornfield-Tukey bridge to evaluate whether the results of that study are relevant to certain other situations, so the critical reader of a case study must examine whether an infer-

ential bridge can be built between this case and other cases of interest to the reader.

Controversy over Method: Experimental Versus Correlational

One of the best known examples of a controversy over method was explicated by Cronbach (1957) in his now classic paper, "The Two Disciplines of Scientific Psychology." Cronbach observed that the field of psychology had divided early into two major streams—the correlational and the experimental. Both these streams share what Cronbach calls the "job of science," which is to ask questions of nature. A discipline, he observes is a method of asking questions and testing answers to determine whether those answers are sound. Correlational psychology is *not* a form of research that uses only one statistical technique—namely, correlation. Those researchers who are deemed correlationists are interested in studying nature as it is, in studying the natural correlations occuring in nature. They are committed to understanding the functional relationships between variations in one set of events or characteristics and variations in another. Thus, they may ask about the relationship between income and achievement, or between the number of physicians per thousand population and infant mortality, or between phases of the moon and the behavior of tides on earth. They see nature presenting itself for inspection and the role of the scientist that of identifying which of the variations that nature presents are associated with other processes or outcomes.

In contrast, experimentalists are interested, as Cronbach observes, only in the variation they themselves create. The experimental method is one where scientists change conditions in order to observe the consequences of those changes. They are interested in understanding how nature is put together, not through inspecting nature as it is, but through introducing modifications or changes in nature in order to better understand the consequences of those changes for sub-

sequent states. They argue that only through the systematic study of planned modifications can we distinguish causal relationships between events or characteristics from mere chance co-occurences. Thus, for example, foot size and vocabulary are correlated in the general population, but that does not mean that large feet cause larger word knowledge (or vice versa). It merely reflects the larger vocabulary size of older (hence, bigger) people relative to children.

All too frequently ignored is the intersection of research methods with the underlying theoretical, political or social purposes of the research being conducted. As I indicated earlier in this paper, research methods are not merely different ways of achieving the same end. They carry with them different ways of asking questions and often different commitments to educational and social ideologies. We can observe this intersection of ideology and method in considering the historical roots of correlational and experimental approaches.

In the scientific world of later 19th century England, the work of Charles Darwin on the origin of the species commanded special attention. Central to his evolutionary theory was the principle of natural selection—nature selects those species or subspecies for ultimate survival that are best adapted to the conditions confronting them. "Survival of the fittest" is a phrase used to describe the process by which individuals and species adapt to variations in environmental conditions in order to survive. The "struggle for life" favors those whose structure and behavior are adaptive to the challenges of their environment and are thus more likely to produce offspring who flourish.

This view of human evolution as a struggle for survival had a substantial impact on prevailing views of society. Buttressed by the centrality of competition and the free market to the economic thinking of 19th century England, a movement called "Social Darwinism" developed. Social Darwinists viewed members of a society as struggling for rewards and undergoing "selection" based on their talents or merits.

Francis Galton, a cousin of Darwin, observed that it was important to study those variations in human abilities and performance contributing most significantly to successful adaptation. He thus began systematically studying those human attributes contributing most to social effectiveness. He assumed those characteristics were enduring traits, unlikely to undergo change. His research was broad indeed, ranging from studies of what he viewed as hereditary genius to investigation of the efficacy of prayer. His research was characteristic of what we now call correlational studies. He developed early forms of the statistical methods which currently underlie correlational research.

Galton's work is historically linked to the brand of social theory that came to be known as Conservative Darwinism (Cremin, 1961). Conservative Darwinists attempted to develop better means for identifying those members of the society who were most likely to adapt successfully and to provide opportunities to those individuals, whatever their social class or family background, to receive education and other perquisites from the society. They constituted the forerunners of the modern testing movement, which can be seen as a way of applying correlational psychology to the problem of identifying the fittest in the society and thereby providing them opportunities for social mobility and leadership.

The testing movement thus began as an attempt to divorce the ability of individuals from their social backgrounds through basing economic and social mobility on performance rather than on patrimony. Ironically, those who now oppose the testing movement base their opposition on the argument that tests merely support and amplify existing social class and ethnic differences.

Opposition to this application of Darwinism to social research developed quickly. Scientists and social reformers questioned the assumption that existing individual or group differences were durable or necessary *by nature*. Indeed, they claimed that such differences were typically historical or social artifacts created by political inequalities. The role of the educator, they asserted, was not merely to develop better ways of identifying the variations already occurring in nature in order to select individuals who are most competent. Indeed, the responsibility of the educator was to identify those interventions in nature that would lead to more successful adaptation and survival for the largest number of human beings. Thus, while survival of the fittest remained the watchword, the responsibility of the educator was to increase the proportion of individuals in the world who are fit, and the responsibility of the educational researcher was to experiment with alternative methods of rendering individuals more fit, more adaptable, than they might otherwise have become. This group was known historically as Reform Darwinists, and their political philosophy is implicit in many applications of experimental methods to educational research.

The goal of the correlationist thus became to understand and exploit the natural and, presumably, enduring variations among individuals, while that of the experimentalist was to create conditions to reduce those variations.

This example of how the two major streams in scientific psychology are ultimately rooted in distinctive political or social commitments is *not* meant to leave you with the impression that these two alternatives must always remain sharply contrasted and never integrated. Indeed, many researchers have devoted their careers to identifying research methods capable of transcending the contrast between experimental and correlational methods. That is a topic, however, that goes far beyond the proper subject of this discussion.

Thus, although Hippocrates was correlating and Galileo experimenting centuries before Darwin, these two strategies of research took on distinctly new ideological implications in the hands of competing Darwinists. In our day, the values commitment implicit in the choice of method is

often unrecognized, even by the investigators themselves. This makes it even more dangerous to treat methodological issues without an understanding or concern for the specific substantive questions being asked. One of the enduring problems in research methodology has been the tendency to treat selection of method as primarily a technical question not associated with the underlying theoretical or substantive rationale of the research to be conducted.

Selecting the method most appropriate for a particular disciplined inquiry is one of the most important, and difficult, responsibilities of a researcher. The choice requires an act of judgment, grounded both in knowledge of methodology and the substantive area of the investigation.

Quantitative and Qualitative Methods

In looking at the differences between quantitative research methods and those typically dubbed qualitative, such as case study or ethnography, we find another type of political or social contrast that is of interest. Quantitative methods, whether correlational or experimental, require large and approximately random samples of individuals. Quantitative approaches require that sampling of both individuals and situations be conducted in order to maximize the generalizability of the findings to the widest possible population. In so doing they tend to sample from individuals and settings *as they are,* rather than *as they might be,* though this may be an overstatement for the experimental approach.

In contrast, it is intriguing to examine the types of setting frequently studied by qualitative researchers. For example, studies of open classrooms, free schools, or other radical, educational innovations are often conducted using case studies or ethnographic methods. In these studies the researcher is attempting to portray the workings of circumstances that differ dramatically from what

typically presents itself in the "natural" functioning of our society and our educational systems. It is as if the researcher is attempting to document with vivid characterizations that nature need not be the way it typically is. The researcher is attempting to communicate that we can create settings far different from those we may discover through random sampling. Moreover, those settings can be both sensible and rule governed.

Often, qualitative researchers studying unusual educational settings accuse quantitative researchers attempting to characterize education, more generally, as committed to maintaining the educational status quo. Qualitative researchers, in contrast, are often committed to demonstrating the viability of truly alternative educational approaches.

I do not want to create these contrasts too starkly. Obviously, many studies of broader educational questions are conducted using qualitative methods, such as some of the more striking investigations of school desegregation, or evaluations of special programs. Conversely, many experimental studies of educational change, such as those conducted in the National Follow-Through experiment, are attempts to introduce significant new approaches to the practices of contemporary education. Here again, however, I have been trying to draw attention to the intricate ways in which the multiplicity of methods we have available in educational research present us, not merely with an enormous technical challenge, but rather with the opportunity to investigate an impressive variety of questions from a rich set of alternative social and political perspectives.

Choosing Among Methods

It is interesting that the most frequently employed educational research methods, and therefore those with the greatest current respectability, are the quantitative methods of experimental,

correlational, quasi-experimental, and survey research. Their disciplinary roots are in agriculture, genetics and other studies of heredity, and actuarial studies of life expectancies conducted two centuries ago in the service of insurance companies. They not only share fairly long traditions in education, but also carry with them the prestige of quantifiable precision. Through the application of modern statistical methods, researchers can more precisely estimate the likelihood and size of errors in estimates of the state of nature, than is usually possible in approaches deriving from anthropology, history, or philosophy. Should we tend to use the more traditional methods because we understand them better and they have a longer track record? John Stuart Mill argued:

If there are some subjects on which the results obtained have finally received the unanimous assent of all who have attended to the proof, and others which . . . have never succeeded in establishing any considerable body of truths, so as to be beyond denial or doubt; it is by generalizing the methods successfully followed by the former enquiries and adapting them to the latter, that we may hope to remove this blot on the face of science.

Yet, an equally brilliant British philosopher of the next century, Alfred North Whitehead, was far less certain that well-developed and understood methods were always likely to be superior. He observed:

Some of the major disasters of mankind have been produced by the narrowness of men with a good methodology . . . to set limits to speculation is treason to the future.

If we are not always well advised to choose the methods that have been used the longest and that we understand best, what of choosing methods on the grounds of precision, on the grounds that some methods provide us a much better base for knowing exactly how much we know and how much is likely to be error? Here again, we are advised to focus first on our problem and its characteristics be-

fore we rush to select the appropriate method. We can again hark back to Aristotle, who made this famous point about precision in the *Ethics:*

Our discussion will be adequate if it has as much clearness as the subject matter admits of, for precision is not to be sought for alike in all discussion, any more than in all the products of crafts. . . . For it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician scientific proofs. (*Ethics,* I: 3) (McKeon, pp. 309-310)

We must avoid becoming educational researchers slavishly committed to some particular method. The image of the little boy who has just received a hammer for a birthday present and suddenly finds that the entire world looks to him like a variety of nails, is too painfully familiar to be tolerated. We must first understand our problem, and decide what questions we are asking, then select the mode of *disciplined inquiry* most appropriate to those questions. If the proper methods are highly quantitative and objective, fine. If they are more subjective or qualitative, we can use them responsibly as well.

The anthropologist Geertz (1973) probably put it best when he said,

I have never been impressed by the argument that, as complete objectivity is impossible in these matters (as, of course, it is) one might as well let one's sentiments run loose. As Robert Solow has remarked, that is like saying that as a perfectly aseptic environment is impossible, one might as well conduct surgery in a sewer. (p. 30)

Geertz also observed, as cited in Wolcott's (1979) discussion of ethnographic research method, "You don't have to know everything to understand something" (Alternative Methodologies in Educational Research Tape Series, L2).

Summary

Let me summarize the important points I have tried to make in this introductory discussion of re-search methodology. What distinguishes research from other forms of human discourse is the application of research methods. When we conduct educational research we make the claim that there *is* method to our madness. Educational research methods are forms of disciplined inquiry. They are disciplined in that they follow sets of rules and principles for pursuing investigations. They are also disciplined in another sense. They have emerged from underlying social or natural science disciplines which have well-developed canons of discovery and verification for making and testing truth claims in their fields. Education itself is not a discipline, but rather a field of study on which we bring to bear the various forms of disciplined inquiry which we have been discussing.

Each of these forms of inquiry asks different questions or has different ways of asking educational research questions. I have tried to illustrate some of the questions characteristic of major forms of educational research methodology. I have also tried to indicate the ways in which the selection of research method is frequently related to theoretical or ideological commitments of the investigator.[1] (Parenthetically, the possibilities of doing certain kinds of social research change as the political and social mood of a society evolves. For example, the notion of randomly assigning individuals to contrasting experimental treatment groups may seem far less acceptable a research strategy in these days of legislation requiring informed consent and protection of human subjects. Can we continue to practice experimental social and educational research and still abide by the law of the land that requires informed consent of all participants in research?)

Finally, each of the examples of research methodology discussed must in some fashion deal with questions of precision and generalizability, although the standards and criteria for these will vary from one form of disciplined inquiry to another.

The neophyte educational re-searcher, when confronted with this imposing array of alternative research methodologies, may be tempted to throw up his or her hands in despair and say, "What can I possibly do to become competent in this field?" I can suggest several answers. First, attempt to become skilled and experienced in at least two forms of research methodology. Facility in only one strikes me as somewhat dangerous, the equivalent of a methodological "Johnny One-Note." Second, be fully aware of the full, rich variety of methods that comprise the family of disciplined inquiry in educational research. Recognize that the most effective programs of educational research are likely to be characterized by what Merton (1975), the distinguished sociologist, or Schwab (1969), the eminent philosopher of education, have called applications of "disciplined eclectic." The best research programs will reflect intelligent deployment of a diversity of research methods applied to their appropriate research questions. Finally, do not limit your education to methodology alone, for only in combining substantive knowledge and methodological competence will you become a well-rounded, effective educational researcher. Here, once again, an insight of Aristotle's is relevant.

Now each man judges well the things he knows, and of these he is a good judge. And so the man who has been educated in a subject is a good judge of that subject, and the man who has received an all-round education is a good judge in general. (*Ethics,* I:3) (McKeon, p. 310)

Remember, selection of appropriate methods is an act of *judgment.*

A variety of methods comprise educational research:historical, philosophical, case studies, ethnographic field studies, experiments, quasi-experiments, surveys. Each is demanding and rigorous and follows disciplined rules or procedures. Taken together these approaches build a methodological mosaic that is the most exciting current field of applied social research—the study of education.

ence. University of California. Berkeley. CA 94720. Dues: $5.00.

SIG: Research Utilization

Purpose: To provide a focus within AERA for activities intended to foster more effective and appropriate application of research results in educational practice.

Contact: Diane McIntyre. 631 Handley Trail. Redwood City. CA 94062. Dues: $3.00.

SIG: Rural Education

Purpose: To encourage educational research relative to rural schools and people in rural America; to provide a forum for the dissemination of findings of such research.

Contact: Everett Edington. ERIC/CRESS. Box 3AP, New Mexico State University. Las Cruces. NM 88003. Dues: $3.00.

SIG: Social Indicators Research*

Purpose: To promote research into the development and use of social indicators in education. including development of theoretical frameworks for social indicators and accessing and evaluating available time-series data.

Contacts: Robert Rossi/Kevin Gilmartin. American Institutes for Research. P.O. Box 1113. Palo Alto, CA 94302. Dues: $5.00.

SIG: Socialization to Adulthood

Purpose: To bring together researchers concerned with the forces that shape healthy socialization to adulthood. Members are also invited to subscribe to *Character*. a periodical about these matters.

Contact: Edward A. Wynne. College of Education. University of Illinois. Box 4348. Chicago. IL 60680. Dues: $3.00.

SIG: Special Education Research

Purpose: To explore areas of educational research of interest to special and general educators. Research areas may include: individual differences. alternative teaching strategies and learning styles.

Contact: Susan J. Schenck. 9 College Way. Education Department. College of Charleston. Charleston. SC 29401. Dues: $5.00.

SIG: State and Regional Educational Research Associations

Purpose: To strengthen state and regional educational research associations by providing a communications network among them and a forum for them.

Contact: Walter M. Mathews. School of Education. University of Mississippi. University. MS

38677. Dues: $25.00 per organization.

SIG: State Education Agencies

Purpose: To focus on research within state education agencies. as well as research involving state and local agencies; to provide an opportunity for members to communicate with each other.

Contact: Nancy Scott. Wisconsin Department of Public Instruction. 125 South Webster. Madison. WI 53702: Dues: $3.00.

SIG: Structural Learning and Instructional Design*

Purpose: To improve understanding of learning and instructional systems and their application in designing effective instruction.

Contact: Joseph M. Scandura. Director. Interdisciplinary Studies in Structural Learning and Instructional Science. 3700 Walnut Street. University of Pennsylvania. Philadelphia. PA 19104. Dues: $5.00.

SIG: Survey Research in Education

Purpose: To provide a forum for researchers interested in improving survey research as a data collection method in educational research.

Contact: Marsha A. Niebuhr. Anderson & Berdie Associates. Inc.. Griggs-Midway Building. 1821 University Avenue. St. Paul. MN 55104. Dues: $4.00

SIG: Systems Research

Purpose: To promote the concept. understanding. development. and application of systems research for planning. problem solving. and evaluation in education and related fields.

Contact: T. A. Ryan. College of Criminal Justice. University of South Carolina. Columbia. SC 29208. Dues: $3.00 AERA Members: $4.00 Non-members.

SIG: Teacher Preparation Curriculum

Purpose: To facilitate the exchange of information and to encourage research into teacher preparation curricula. including the study of teacher behavior. training techniques. and research methodologies.

Contact: Katherine O'Donnell. Bank Street College. 610 W. 112th Street. New York. NY 10025. Dues: $2.00.

SIG: Text Design and Learner Strategies

Purpose: To facilitate communication among researchers. publishers. writers. and teachers who are interested in improving text design and examining learner strategies.

Contact: Shawn M. Glynn. Department of Educational Psychology. University of Georgia. Athens. GA 30602. Dues: $2.00.

SIG: Training in Business and Industry

Purpose: To foster communication among educational researchers. evaluators. instructors. technologists. course developers. and performance analysts in business and industry concerned with the adult learner in training environments.

Contact: David Kerr. Sterling Institute. 700 Dodge Center. Georgetown. 1010 Wisconsin Avenue. NW. Washington. D.C. 20007. Dues: $7.00.

SIG: Urban Education

Purpose: To provide structure within AERA for research and evaluation that is concerned with. and focuses on. education in the urban environment.

Contact: David E. Kapel. School of Education. University of Louisville. Louisville. KY 40292. Dues: $3.00.

SIG: Vocational Education

Purpose: To foster interchange of ideas and methods between the general and vocational educational researcher. and aid in design and conduct of research related to vocational-technical education programs.

Contact: John D. Skinkle. Interdisciplinary Education. College of Education. Texas A&M University. College Station. TX 77843. Dues: $15.00.

SIGs In Formation—

Families as Educators

Purpose: To promote the study and dissemination of information on family social processes and home-school relationships that support children's education and development. *Contact:* Diane Scott-Jones. LRDC. University of Pittsburgh. Pittsburgh. PA 15260. *Dues:* $2.00—AERA members: $5.00—Nonmembers.

Research on the Teaching and Learning of Economic Education

Purpose: To disseminate research finds on the teaching and learning of economics. K-adult *and* to strengthen the disciplinary ties between education research and research on economics education. *Contact:* John Sumansky. Joint Council on Economic Education. 1212 Avenue of the Americas. New York. NY 10036. *Dues:* $2.00.

Early Adolescence

Purpose: To provide a forum for research on the education of youngsters aged 10-15 and also to encourage an interdisciplinary perspective in research and attempt to stimulate communication between researchers and practitioners. *Contact:* Bruce Mackenzie-Haslam. School Management & Organization Studies. Mail Stop 16. NIE. Washington. DC 20208. *Dues:* $5.00.

SHULMAN

Note

[1] For a particularly instructive example of this principle in the history of education, see Tyack (1976).

References

Aristotle, *De anima (On the soul)*. In R. McKeon (Ed.), *Introduction to Aristotle*. New York: Modern Library, 1947.

Aristotle, *Nicomachean ethics*. In R. McKeon (Ed.), *Introduction to Aristotle*. New York: Modern Library, 1947.

Cornfield, J., & Tukey, J. W. Average values of mean squares in factorials. *Annals of Mathematical Statistics*, 1956, *27*, 907-959.

Cremin, L. A. *The transformation of the school*. New York: Vintage Books, 1961.

Cronbach, L. J. The two disciplines of scientific psychology. *American Psychologist*, 1957, *12*, 671-684.

Cronbach, L. J., & Suppes, P. (Eds.). *Research for tomorrow's schools: Disciplined inquiry for education*. New York: MacMillan, 1969.

Geertz, C. Thick description. In C. Geertz (Ed.), *The interpretation of cultures*. New York: Basic Books, 1973.

Merton, R. K. Structural analysis in sociology. In P. Blau (Ed.), *Approaches to the study of social structure*. New York: Free Press, 1975.

Schwab, J. J. The practical: A language for curriculum. *School Review*, 1969, *78*, 1-23. (Reprinted in Schwab, J. J. *Science, curriculum and liberal education*. Chicago: University of Chicago Press, 1978. [pp. 287-321]).

Tyack, D. Ways of seeing: An essay on the history of compulsory schooling. *Harvard Educational Review*, 1976, *46*, 355-389.

Lee S. Shulman is Professor of Educational Psychology and Medical Education; Codirector, Institute for Research on Teaching. College of Education, Michigan State University, East Lansing, MI 48824. Specializations: Research on Teaching; Psychology of Judgement, Decision Making and Problem Solving.