

RIGOR VERSUS VIGOR: SOME DUBIOUS ISSUES IN THE DEBATE ON RESEARCH PHILOSOPHY¹

HERBERT C. KELMAN

The University of Michigan, U.S.A.

Social and behavioral scientists devote a considerable amount of their time and energy to discussions about alternative philosophies of research. Such discussions are especially prominent in interdisciplinary research settings, where they often take the form of debates between the disciplines. For example, during the past fifteen years, research in mental hospitals has drawn on the contributions of anthropologists, psychiatrists, psychologists, and sociologists. Investigators coming out of these different research traditions have tended to formulate the research problems in different ways, to focus on different units of analysis, and to favor different research methods. These differences tend, on occasion, to take on ideological overtones and to degenerate into contests over the relative virtues of the different disciplines and over their relative faithfulness to the precepts of science or to the phenomena under study.

Debates about alternative philosophies of research may also take place within a given discipline. Thus, for example, within the field of political science, a debate between "behavioralists" and "traditionalists" has been going on at least since the end of World War II. The behavioralists favor quantitative research and draw their concepts from such fields as psychology and sociology. The traditionalists are more inclined to follow historical, descriptive, and normative approaches to their subject matter. Again the debate has often tended to be ideological, with each side questioning the very legitimacy of the enterprise in which the other side is engaged.

Within the field of social psychology, which is interdisciplinary by its nature—being rooted in psychology and sociology, and influenced by other disciplines as well—it is inevitable that differences in research philosophy would arise. These differences are in part related to the disciplinary origins—i.e., psychology or sociology—of those who call themselves social psychologists. Many of the differences, however, cut across these disciplinary lines and arise as much within the two subsets of social psychologists as they do between them.

Social psychologists differ in their view of the proper *level of analysis* for

¹ This paper will appear as a chapter in a book on *The Value Context of Social Research*, to be published by Jossey-Bass, Inc. (San Francisco) in the fall of 1967. It is based, in part, on a paper that was originally prepared while the author was working at the National Institute of Mental Health (U.S.A.).

social-psychological research. Thus, some investigators (including some who were originally trained in sociology) regard social psychology as essentially an extension of general psychology. They attempt to analyze social behavior in terms of general psychological principles, operating under the special conditions provided by the social situation. Their strategy is to take as their starting-point models developed for the analysis of simple situations and to see how far these can be pushed in the analysis of the more complex situations in which social behavior occurs. Other investigators (including many who were originally trained in psychology) assume that social behavior cannot be effectively reduced to the level of individual behavior and that social psychology must, therefore, develop its own level of analysis. Their strategy, accordingly, is to start out with models that represent—as parsimoniously as possible—the complexity of the phenomena that they hope to illuminate.

Social psychologists also differ in the *research methods* that they prefer to use. Roughly speaking, one might distinguish between those whose orientation is primarily quantitative and those whose orientation is more equalitative—or at least clinical and non-statistical.² Within each of these two sets of orientations, one can make further distinctions, which often reflect important differences in research philosophy. Thus, among the "statistically" oriented social psychologists, there are those who rely primarily on the experimental method, involving the active manipulation of variables to be studied and observation of their effects; and those who rely primarily on the use of survey methods to obtain opinion data and behavior reports from selected samples of the population. Similarly, among the more "clinically" oriented social psychologists, one can distinguish between those who prefer to use structured methods to obtain their data on individuals and groups, such as personality tests, structured interviews, and systematic group observation; and those who prefer to use the methods of participant observation and community study in the anthropological tradition.

As in the case of other differences in research philosophy that I have touched upon, the differences among social psychologists are sometimes ideologized. Proponents of one orientation may insist that they alone are truly scientific; proponents of another orientation may counter that they alone are truly relevant. Debates about research philosophy are a healthy and necessary part of any field of investigation, but when they take the form of arguments about the general virtue of "my" concepts and methods as compared to "your" concepts

² These terms are only approximations. I would include in the second category studies in the factor-analytic tradition that collect large amounts of data on a single individual or a small number of individuals. These studies are certainly quantitative, in the sense that they work with numerical scores and subject these to various kinds of mathematical operations. They are, however, non-statistical, in the sense that they are not primarily concerned with drawing inferences on the basis of sampling procedures (whether these be random sampling from a surveyed population or random assignment of subjects to experimental and control groups).

and methods—without regard to the specific uses to which these are put—then they are bound to be fruitless. More often than not, when we argue about the relative merits of psychological vs. sociological, or quantitative vs. qualitative, or experimental vs. naturalistic approaches, we are dealing with false issues.

Sometimes the issues are false because they are based on a misconception of what the "opposing" approach or its particular representative stands for. Thus, antagonists in the debate between psychological vs. sociological orientations may be operating with stereotyped and anachronistic notions of how the other discipline approaches its problems. I have heard psychological analysis criticized because, supposedly, it seeks to explain group or collective behavior in terms of the idiosyncratic personality characteristics of individuals. I have heard sociological analysis criticized because, supposedly, it is based on armchair speculation or at best on casual observation. Such conceptions may, perhaps, have been accurate at one time, and they may still be good characterizations of how some psychologists or some sociologists go about their respective tasks. But they are highly distorted pictures of the dominant trends in these disciplines, and one can certainly not assume that they are accurate characterizations of the approach of any *given* psychologist or sociologist. Above all, such stereotypes ignore the fact that within each of the two disciplines, there is wide diversity in the concepts, methods, and levels of analysis used.

Similar misconceptions may be at the root of certain other dubious controversies. For example, some may be critical of a qualitative approach because they see it as an undisciplined selection of illustrations that can support any point the investigator wishes to prove. While we have all seen research that conforms to these unhappy stereotypes, they are clearly caricatures and perversions of the skilled and intelligent use of quantitative or qualitative analysis. Whenever debaters operate on the basis of such mutual misconceptions, there is no possibility for a true confrontation and resolution of issues. The antagonists have not even come to the point of discovering whether a real difference exists between them.

Even when two antagonists are genuinely talking to each other and real differences in research approach do exist between them, they may be debating false issues. This happens when the debate contrasts two alternative approaches that are not really alternatives at all. The debate assumes that they are two different ways of doing the same thing, when in fact they turn out to be two different ways of doing two different things. It makes no sense to argue whether one set of concepts or methods is better than another when the two are designed to deal with different kinds of problems. The question must always be: better for what? Two investigators—both of whom may, for example, be studying social interaction—may be completely justified in working with different variables and units of analysis, if one is primarily concerned with the

effects of interaction on individual personality and the other with its effects on the social system. Similarly, the investigator concerned with defining the contours of a phenomenon cannot reasonably be criticized for his failure to use an experimental approach, just as the investigator concerned with establishing functional relationships between variables cannot be criticized for his failure to use a more naturalistic approach. Debates that ignore the purposes that the competing approaches are designed to serve cannot provide much illumination.

The real issue is whether a particular research approach is appropriate to the questions that the investigator is trying to answer. Beyond that, there is the more complicated matter of whether these questions are worth asking at all. This, of course, raises the larger issues of the definition of the social researcher's task and the nature of the enterprise in which he is engaged. These real issues—concerned with the fit between our research approaches and our research problems, and with the nature of our essential task—form the background for much of the debate on research philosophy that is taking place. They tend to be masked, however, and relegated to the periphery, when the discussion focuses on false dichotomies, rooted in competing disciplinary loyalties and scientific models.

Many of the dubious issues that I have already mentioned turn on one or another aspect of what might be called the "rigor vs. vigor" controversy. In the pages that follow, I shall try to spell out why the controversy on this level is of questionable validity, and how some of the real issues in the evaluation of social research cut across this particular dichotomy. I shall then turn to one special type of research approach, which is generally seen as the model of rigor—namely the experimental method. By examining experimental work in the context of what I regard as the essential task of the social scientist, I hope to demonstrate that, even in this case, the rigor vs. vigor formulation provides a misleading picture.

The Rigor Versus Vigor Controversy

Some years ago, Kenneth Boulding took certain liberties with the lyrics of W. S. Gilbert to drawn portraits of the scientist and the humanist. With a few further modifications, Boulding's lyrics can serve to demonstrate some of the differences between the proponents of rigor and the proponents of vigor. With apologies to both Gilbert and Boulding, then, let me present this brief debate between representatives of the two schools of thought:

R. Conceive me if you can
A rigorous young man,
A physical-causative,
Logical-positive,

White-coat-and-rat young man,
 Who has decided bent
 Towards experiment,
 And Oh! what a wrath is his
 If a hypothesis
 Claims it's self-evident.

- V. Conceive me if you can
 A vigorous young man,
 A most analytical,
 If not always critical,
 Raw-stuff-of-life young man,
 Who is persuaded that
Man differs from the rat.

To pierce men's reality
 He views their totality
 In their native habitat.

- R. A serious-aims young man,
 A Theory-of-Games young man,
 A very stochastic
 Iconoclastical,
 Testable-claims young man.

- V. A very unique young man,
 A reach-for-the-peak young man,
 A flagrantly mystical
 Most unstatistical
 Rather oblique young man.

- R. A one-way-screen young man,
 A keep-it-clean young man,
 A fine-observational,
 Most operational,
 Truth-is-what's-seen young man.

- V. A truly all-round young man,
 Yet very profound young man,
 An I'll-only-do-it-if-
 It-is-intuitive
 Truth-at-a-bound young man.

- R. A look-at-the-fact young man,
 A get-out-and-act young man,
 A set-up-a-project-

Where-money's-no-object,
Foundational-backed young man.

V. A very arm-chair young man,
A no-questionnaire young man,
A find-it-out-many-ways,
Know-it-all-anyways,
Devil-may-care young man.

R. A facts-by-the-yard young man,
A punchable-card young man,
Statistical-tabular,
Special-vocabular,
Work-very-hard young man.

V. A fitful-gleam young man,
A beautiful-dream young man,
A most metaphorical
Highly rhetorical
Letting-off-steam young man.

Needless to say, these portraits are stereotyped caricatures of the two sides, but they do transmit some of the flavor of the debate between them. The proponents of rigor stress the importance of obtaining hard facts through the use of exact methods. They prefer research situations that are maximally structured and, if at all possible, based on the experimental manipulation of the independent variables; research instruments that are impersonal and do not require too much filtering through the mind of the investigator; and dependent variables that can be stated in terms of quantitative indices and subjected to statistical analysis. Objectivity, precision, and replicability are, in their view, central requirements of a scientific analysis. The proponents of vigor, on the other hand, stress the importance of research methods that will capture the real-life flavor of the phenomenon under study, in all of its fullness and richness. They prefer to make their observations in natural settings, free from the manipulations of the investigators; to use themselves as active instruments for sorting, understanding, and integrating what they observe; and to present their findings in ways that properly reflect the totality of man and society and the complexity of social behavior. Precision of methods is, in their view, less crucial than relevance to the questions that social science proposes to answer and truthfulness to the phenomena of social life.

The different components of the two syndromes I have described do not necessarily always go together, nor do most social psychologists clearly identify themselves with one or the other of these syndromes. Yet this division is at

at least an undertone, in much of the debate and position-taking on questions of research philosophy and of the most productive direction for the future development of the field.

The question of whether experimental or naturalistic approaches are ultimately superior is meaningless, because it cannot possibly be answered in general terms. The choice of appropriate methods depends, first of all, on the nature of the problem under investigation—the kinds of questions that the investigator is asking. A fetishistic insistence on rigorous methods is bound to close off many promising sources of insight and information. Conversely, a squeamish reluctance to pin things down lest the sanctity of the phenomenon be destroyed is likely to keep us from ever attaining reliable knowledge. Depending on our purposes, a sacrifice in either precision or naturalness may be completely justified. For example, if we are interested in identifying the variables that define a particular phenomenon and in tracing its course of development, then we have to observe it in its natural setting. A less rigorous and more impressionistic approach is clearly called for here, since it yields the data we need and cannot obtain in any other way. On the other hand, if we are interested in establishing functional relationships between two sets of variables, then there is no substitute for an experimental approach, even though this forces us to rely on artificial situations of limited generalizability.

The choice between experimental and naturalistic research depends also on the stage of development of our research area and of the specific problem on which we are working. With respect to the research area in general, the "scientific ethos" requires us to use the best methods available for dealing with the questions we have posed. The unavailability of precise methods is certainly no reason for turning away from these questions. We tackle them as best we can, noting the limitations of our methods and gradually working toward their improvement.

With respect to the stage of development of the specific problem, different approaches are called for at different points in time. In the initial stages of work on the problem, when the investigator is interested in gaining an intuitive understanding of the phenomenon and in developing hypotheses, a clinical-impressionistic approach is most appropriate. When he is ready to test these hypotheses in the form of functional relationships, an experimental approach or some approximation thereof is usually called for. In a later stage, when the investigator is interested in testing the generality of his hypotheses in real-life settings, he may again turn to naturalistic observations or to the use of survey methods.

Finally, the choice of experimental or naturalistic approach must also depend on the preferred style of the investigator. Some are more comfortable with one approach and some with another—for reasons of personality, train-

ing, or esthetic preference. When an investigator uses the approach that is most congenial to him, he is likely—other things being equal, of course—to be more creative and to make a better contribution. Thus, what is a productive method for one investigator is not necessarily so for another. Each investigator should feel completely free to select his own preferred style—without, however, claiming that it is therefore objectively better for all problems. He should keep in mind the limitations of the particular approach he has selected, apply it only to those problems to which it is naturally suited, and derive from it only those conclusions that it is capable of producing. In short, if he writes musical comedy, he should not try to pass it off as epic drama.

Two aspects of the rigor vs. vigor issue deserve special comment. One is the question of quantitative vs. qualitative analysis; the other the question of a holistic vs. an elementalistic approach.

Qualitative analysis is sometimes criticized on the grounds that the investigator has no conception of what he has really found and is highly susceptible to the fallacy of the positive instance. Quantitative analysis, on the other hand, may be criticized for the investigator's tendency to let the true phenomenon pass him by while he is simple-mindedly and obsessively counting the irrelevant. Thus, critics of a qualitative approach argue that the conclusions derived from it may be interesting, but are probably not true; while critics of a quantitative approach argue that the conclusions derived from the latter may be true, but are probably not interesting.

The mere posing of the issue in terms of qualitative vs. quantitative is bound to be misleading. In a very basic sense, all scientific statements are inevitably quantitative, although the counting and measurement may be only implicit. Even a case study, for example, makes some implicitly quantitative statements: When we describe certain patterns of behavior that are characteristic of the person or community studied, then we are essentially saying that this pattern occurs *frequently* and in *many* kinds of situations. Often, moreover, we select a particular case for investigation because it is deviant or extreme—in other words, because it possesses a particular trait or set of traits to a *greater* or *lesser* extent than other comparable cases.

Similarly, scientific statements generally have a qualitative component, although it too is not always made explicit. Propositions that state the relationship between two quantitative variables apply only if certain background conditions obtain. For example, the nature of the relationship between social pressure and conforming behavior depends on the nature of the situation in which the pressure is exerted, on the cultural context in which the behavior is observed, and on the motivational set that the person brings to the experience. In principle, these background factors can all be stated in quantitative terms, but in practice we rely—and must rely—on qualitative statements of the limit-

ing conditions of our propositions. In empirical work, it is easy to ignore these conditions, because they are built—as constants—into the natural situations we select for study, or the experimental situations we create in the laboratory. Nevertheless, these qualitative factors have an important bearing on our findings and—no matter how strong our predilection for quantitative statements—must be taken into account.

Thus, the crucial issue that cuts across the quantitative-qualitative dichotomy is whether, in any given study, the investigator is collecting the data appropriate to his particular purpose, and whether he is drawing conclusions that his particular data entitle him to draw. Those of us who use quantitative methods must always ask ourselves whether what we are counting and measuring is meaningfully related to the phenomena we are interested in, or whether we are just counting and measuring whatever happens to be readily available. The primitive, simple-minded quantifier, who decides what to measure on the basis of the instruments he has at hand, with little regard to the problems he hopes to illuminate, is certainly deserving of criticism. His culpability is compounded if he mistakes his irrelevant or partial index for the real thing. A hypothetical example of such a misuse of the quantitative approach would be an investigator who is interested in the level and determinants of "religiosity" in modern society, who takes church attendance as his sole index of religiosity because it can be assessed easily and reliably, and who then proceeds to draw broad conclusions about religiosity on the basis of these data. Critics of a quantitative approach would undoubtedly point to a study of this kind as a good illustration of the foolishness of trying to investigate religiosity with quantitative methods. What is wrong with this study, however, is not that the investigator used quantitative methods, but that he based his quantitative indices on the wrong data—or at least on an insufficient range of data—and that the data he did collect did not entitle him to draw the conclusions that he proceeded to draw. These errors are not inherent in the use of quantitative analysis as such.

For qualitative research, too, the basic question is whether the investigator has obtained the data appropriate to the conclusions that he wishes to draw. The primary danger here is not that he will focus on irrelevant data that can be easily measured, but that he will focus entirely on positive instances. Qualitative research does not have as many built-in safeguards against the tendency to find what one is looking for. An investigator eager to illustrate a particular phenomenon may find many cases that are consistent with his analysis, while missing many other cases that might be inconsistent with it. But, again, this fallacy is not inherent in the use of qualitative analysis as such. In qualitative as well as in quantitative analysis, the question is: What kinds of data does the investigator use for his analysis, and does he take the nature of these data into account in drawing his conclusions? Specifically, in evaluating qualitative re-

search, we would ask whether the investigator has deliberately looked for disconfirming cases as much as for confirming ones—whether he has arranged his observations in such a way that contrary findings have an equal opportunity to emerge.

A second aspect of the rigor vs. vigor controversy on which I wish to comment briefly is the apposition between a holistic and an elementalistic approach to human behavior. One component of the "vigor" syndrome (though by no means a universal one) is the insistence on dealing with "the whole man," in contradistinction to segmental, elementalistic approaches which only deal with parts of man and destroy his basic unity. To my mind, this is a false issue, because I do not regard the whole man as the business of social science. He is the proper business of poetry, philosophy, religion, everyday interpersonal relations and, to a certain extent, of psychotherapy. But the task of the social and behavioral sciences is to dissect and analyze the behavior of men and societies, to break it down in terms of theoretical constructs and genotypical formulations, and in this way to increase our systematic understanding of it. It is in the nature of social research to deal with parts of man, artificially separated out from the richness and unity of the total personality. The uniqueness and wholeness of the individual disappears at the hand of the generalizing social scientist, just as the poet's nature disappears at the hand of the natural scientist.

The holistic vs. elementalistic dichotomy is sometimes used to distinguish between psychoanalysis and academic psychology. But certainly psychoanalysis does not deal with the whole man. Among the greatest contributions of psychoanalytic theory are the different ways in which it slices man—and, until recently, the theory has not even concerned itself with some of these slices. It is precisely because of the fact that psychoanalysis is a scientific—hence not a holistic—system, that it is not a proper philosophy of life and cannot substitute for an adequate metaphysics and ethics. Man's ultimate meaning, his place in the universe, the nature of good and evil—all of these must be examined outside the scientific realm.

While social science is, by definition, concerned with parts, the picture of man that emerges from its theory and research must be *consistent* with the nature of man and society as they manifest themselves in their phenotypical wholeness. For example, a theory that cannot encompass such human and societal characteristics as self-sacrifice and love, or power and murder, is obviously of limited value. Another and related issue is whether the units of analysis that are used in a given theory and research program are appropriate to the problem with which they are intended to deal. This, it seems to me, is the real issue behind much of the holistic vs. elementalistic controversy. For example, I would question the use of a stimulus-response model for social psychology, not because it is too elementalistic, but because its units—which may be quite appro-

appropriate for the study of conditioning—are not equally appropriate to the complexities of social interaction. I would favor such concepts as social role and self-presentation—though they too are analytical and segmental—because I regard them as more appropriate to the level of analysis at which the social psychologist operates.

In sum, I have maintained that the rigor vs. vigor controversy and its various sub-controversies—such as those involving the relative virtues of quantitative vs. qualitative or holistic vs. elementalistic approaches—tend to focus on false issues. The real questions—stated in their most general terms—on which the evaluation of a given line of social research hinges, are how *systematic* and how *imaginative it is*. These questions are often masked by the rigor vs. vigor debate, because they sound deceptively similar to that dichotomy, while in fact they cut across it.

To be systematic is not the same as to be rigorous—in the sense of using experimental methods and precise, quantitative measures. Systematic work in social science refers to an organized and disciplined way of thinking about social behavior and of moving back and forth between conceptualization and evidence. Systematicness is not linked to a particular set of methods, although it does imply an awareness of the limitations of the methods one is using and of the conclusions that can properly be drawn from them. It is a relevant criterion in the evaluation of any social research, regardless of its form, and within each research tradition there may be wide variations in how adequately this criterion is met. Thus, qualitative, clinical work may be quite systematic, even though it is not rigorous. Conversely, quantitative, experimental work—though highly rigorous—may be quite unsystematic, if it involves the mere accumulation of empirical data without any attempt to relate these to efforts at conceptualizing social behavior.

The criterion of imaginativeness also cuts across the rigor vs. vigor dichotomy. The tendency to credit the “looser,” clinical approaches with profound insights and to equate the “cold,” statistical approaches with arid recapitulations of the obvious is a romantic notion that does not always conform to reality. The use of naturalistic methods, which attempt to capture the richness and the real-life flavor of the phenomenon under investigation, does not guarantee that the research will be imaginative. Conversely, quantitative and experimental work of the most rigorous kind may well be highly imaginative and imbued with a creative spark.

The real issue in the evaluation of social research revolves, in short, around the thinking and the imagination that it represents. There is no substitute for good thinking and good imagination, no matter what methods we use. And no method automatically insures us of either one of these.

EXPERIMENTAL RESEARCH IN SOCIAL PSYCHOLOGY

In distinguishing between the different purposes of different types of research, I pointed out that naturalistic research is particularly appropriate for exploring the dimensions of a problem and developing hypotheses, while experimental research is particularly appropriate for testing hypotheses about the functional relationships between different variables. This is, essentially, the usual distinction between research undertaken for purposes of discovery and research undertaken for purposes of verification. The latter clearly calls for a greater degree of rigor—highly structured designs, quantitative methods, and statistical analysis. By and large, I would agree with this formulation, but I would argue that—at least in social psychology—the distinction between discovery and verification is not as sharp as we sometimes maintain. In line with this position, I would like to propose a somewhat different perspective for viewing the functions and contributions of experimental research in this field.

The central features of research design in experimental studies are based on the assumption that we are engaged in efforts to verify general propositions. To this end, we are concerned about random assignment of our subjects to experimental conditions, about providing appropriate controls for our crucial comparisons, and about eliminating alternative hypotheses. In short, we set up our studies in such a way that they will allow us to verify propositions according to the usual standards of experimental method. If experimental research is to maximize its potential contributions—which, as I shall point out shortly, I regard as very considerable—then its practitioners must follow these procedures faithfully. The value of this work rests on our playing the experimental game according to its rules. And, as long as we do, I feel that we are entitled to the privilege of talking the experimental game—of using the language of testing and verification adopted from the natural sciences.

Let us not, however, deceive ourselves about the status of what we have found when we complete an experiment in social psychology. I would maintain that the findings of social-psychological experiments—even if they have been replicated a dozen times—can hardly be thought of as experimentally verified and established laws of nature. I assume that it is reasonable to view experimental findings in the natural sciences—and perhaps even in certain areas within psychology—in these terms. In social psychology, however, the gap between the laboratory and the real world is so great that one is hardly justified in the conclusion that what has been established in the laboratory constitutes a verified fact about nature. This is not because the laboratory situation is unreal, but because it has its own reality. The characteristics of the experimental situation in which we put our hypotheses to the test are related in ways that are largely unknown to the characteristics of the situations to which we hope to

generalize our findings. First of all, the laboratory situation as such has unique characteristics of its own, which are only partly understood—although there has recently been a healthy concern with exploring the social psychology of the psychological experiment. In view of the special characteristics of the laboratory situation, it may well be that the whole array of findings based on laboratory studies is applicable only to behavior in the laboratory—and in a subset of other social situations that are similar along certain crucial dimensions. In addition to the unique characteristics of laboratory situations in general, the situation created for any given experiment or series of experiments has special characteristics of its own. As I pointed out earlier, in discussing the role of qualitative factors, in setting up an experiment we devise certain background conditions that remain constant throughout. It may well be that the relationships found hold only in situations that share some of these background conditions. Similarly, in any given experiment we operationalize our independent variable in a particular way and we measure our dependent variable in a particular way. We know that the particular ways in which our variables are operationalized and measured often make a difference in the relationships obtained. It is very difficult, therefore, to have any reasonable assurance that relationships found in the laboratory apply to the wide range of real-life situations, characterized by different background conditions and different manifestations of the independent and dependent variables.

In view of the idiosyncratic and unrepresentative nature of experimental situations in social psychology, and in view of our limited knowledge of the dynamics of those situations, we cannot reasonably equate the confirmation of an experimental hypothesis in the laboratory with the verification of a general principle in nature. Our ability to draw such conclusions may be enhanced as we explore a particular relationship in a wide variety of settings—both experimental and natural, and as we learn about the special characteristics of the laboratory situation. At least as of now, however, our field has certainly not achieved this stage in its development.

Does that mean, then, that our experimental work is really *just a game*, that we are merely playing scientist without contributing anything substantial to an understanding of the general principles of social behavior? My answer is an emphatic no. In my view, experimental research can make enormously important contributions to social-psychological knowledge. These contributions, however, take the form of providing *unique inputs into systematic thinking about social-psychological processes*, rather than of establishing laws about social behavior.

I regard systematic thinking about man and society as the central and essential task of the social scientist. I would not insist that it is his *only* legitimate task. He may also, for example, apply his skills to the solution of certain practi-

cal problems. Or he may devote himself to the collection and processing of various kinds of social data (such as demographic or public opinion data) because of their historical interest or policy relevance, rather than because of their theoretical significance. But systematic thinking—continually confronting all manner of evidence and aiming for the development and refinement of general propositions—is the social scientist's task *par excellence*. It is in this context that I would want to evaluate the contributions of experimental methods, and it is in this context that I see a unique role for the experimental approach. There are at least four ways in which experimental research can feed into the process of thinking about social behavior more effectively than any other type of method:

(1) The requirement to translate our concepts into experimental operations imposes a discipline on our thinking that might otherwise be lacking. As long as we remain at the level of manipulating words, we can gloss over certain conceptual difficulties and avoid the necessity of really resolving certain ambiguities in our thinking. Once we attempt, however, to specify the conditions necessary for testing our propositions, to create laboratory situations encompassing these conditions, and to manipulate our variables through concrete operations, we begin to discover ambiguities that had remained unnoticed and we are forced to face difficulties that we had been avoiding. The necessity of devising an experiment forces us to commit ourselves—to state clearly what our concepts mean and to pin down precisely what relationships we expect.

(2) Experiments offer us an opportunity to observe causal relationships, which can usually be inferred only indirectly and tenuously from other types of evidence. An experiment cannot, of course—as I have already stressed—indicate the generality of the relationship found. It may, however, provide a very important input into our thinking by showing that the causal relationship between two variables *can* be in the particular direction found, at least under certain circumstances. This kind of information is useful in identifying the dimensions on which our conceptual efforts ought to focus and in suggesting lines of thought that are likely to be more or less productive to pursue.

(3) Experiments provide operating models of the social-psychological systems that we are interested in exploring. They allow us to study certain processes in situations that we have deliberately created and that, therefore, have some advantages over the real-life situations to which we ultimately hope to generalize. The experimental situations are simpler, being stripped—insofar as possible—of extraneous variables and historical complications; they are situations whose histories and dimensions are more fully known to the investigator; and they are—at least to a limited extent—subject to the investigator's control. It is thus possible to observe the operation of the specific variables of interest in a detailed and relatively uncontaminated fashion. It is also possible

to extend our range of observations to hypothetical situations that do not exist and have never occurred in real life. Along with its advantages, the stylized and artificial character of the experimental situation also contains certain disadvantages—notably the difficulty in generalization. When taken in conjunction with observations from real life, however, the observations of a hypothetical model in action provide unique inputs into our thinking about social processes and their potentialities for change.

(4) If theoretical thinking is to remain productive, it cannot feed entirely upon itself, but requires periodic stimulation by new inputs from outside sources. For the social scientist, his empirical observations constitute the major source of such inputs. Every type of observation has something unique to offer as a stimulus to new thinking. One of the unique contributions of experimental observations is that they often derive from novel, atypical situations. Another unique contribution is the possibility of accumulating findings from a series of systematically interrelated experiments, which together point a new direction. It is interesting to note that unanticipated experimental findings—which are not too desirable from the point of view of verifying propositions—are particularly useful inputs into new thinking. Unanticipated findings call the investigator's attention to variables he had not thought of before and suggest interpretations and qualifications that he had not considered.

These are some of the special ways, then, in which experimental work feeds into the process of thinking about social behavior and that make the experimental method such an important tool in social research. It is on their embeddedness in this longer conceptual process, rather than on their direct contribution to the body of verified laws about social behavior, that the significance of experimental studies rests. *An experimental finding, at least in our field, cannot very meaningfully stand by itself.* Its contribution to knowledge hinges on the conceptual thinking that has produced it and into which it is subsequently fed back.

There are several implications to the present view that an experimental finding cannot stand by itself. The most obvious implication, with which no theoretically oriented experimenter would disagree, is that *findings from any single experiment* cannot stand by themselves. Contributions are of necessity cumulative; only as a series of experiments—either by the same investigator, or by different investigators working on related problems—build upon each other, can we begin to formulate meaningful conclusions. There are different points of view about what constitutes a good research program, likely to produce a cumulative effect. Some experimenters prefer to narrow in on their problem, using a variety of experimental situations that focus on related issues, in the hope that the nature of the phenomenon they are exploring will gradually become clarified. Others prefer to use a single experimental situation, systematically varying

all the variables that are—for theoretical or empirical reasons—potentially significant, in the hope that they can thus pin down the whole array of factors controlling the phenomenon under investigation. Both types of experimental programs may provide useful inputs into thinking about social behavior, though each has different strengths and weaknesses. Thus, the "narrowing-in" approach provides a better basis for assessing the generality of the phenomenon, while the "varying all the variables" approach is more useful in systematically identifying the dimensions that ought to be considered.

A second implication of the present view, with which some experimental social psychologists may not concur, is that *findings from experimental research* cannot stand by themselves. Because of the gap between laboratory situations and the range of real-life situations to which we want to be relevant, we cannot base our general propositions on the findings from experimental research alone. Our thinking must be informed by data from a wide variety of sources—population surveys and correlational research, participant observation and community studies, analysis of documents and of individual cases, historical studies and ethnographic reports. This is not to say that every investigator must work in all of these traditions, or even in more than one. There is no reason why a social psychologist who is trained in experimental work, is good at it, and enjoys it, should not devote himself to laboratory studies. He should, however, be aware of the limitations of experimental work and to its place within the larger context of systematic thinking about social behavior. Moreover, if he is well trained, he should at least be able to draw on observations from other sources as he contemplates his experiment—both in the process of defining his problem and developing his hypotheses before he begins his experiment, and in the process of checking the generality of his findings and exploring what they mean in action after the experiment is over.

Finally, the present view implies that *empirical facts* in our field—especially, but not exclusively, those based on experimental research—cannot stand by themselves. It is not the facts that constitute a contribution, but what is done with them. It may, of course, happen that an experimenter does little or nothing with his findings himself, but they become the raw material for the conceptual work of others. In any event, the ultimate value of experimental findings depends on the quality of the thinking in which they are embedded. It follows that a study that is procedurally clean and well-designed, but unrelated to a serious conceptual problem, is less valuable than a messier study that forms part of a systematic process of thinking about social behavior. In other words, even though rigor is one of the unique and central contributions of experimental method, rigor as an end in itself is self-defeating. In the final analysis, no method can substitute for intelligence and imagination.

ABSTRACT

Many dubious issues turn on one or another aspect of what might be called the "rigor vs. vigor" controversy. They are generally "false issues" of questionable validity, such as naturalistic vs. experimental, psychological vs. sociological, quantitative vs. qualitative, and holistic or molar vs. elementalistic or reductive approaches.

The real issue is whether a particular approach is appropriate to the question the investigator is trying to answer, and the even more complicated matter of whether the question is worth asking.

Rigor, the stressing of hard facts through the use of exact methods, is contrasted with vigor, stressing methods that will capture real-life flavor in all its fullness, and this dichotomy is used to illustrate that the choice of a proper method depends on the nature of the problem under investigation, the stage of development of the research area, the specific problem on which the work is being done, and the preferred style of the investigator. The value of the research depends upon how systematic and imaginative the approach is, for no experimental finding in our field can stand meaningfully by itself, apart from the conceptual thinking and related theory, or apart from cumulative supportive evidence.

RESUMEN

En la controversia del "rigor" versus "vigor," se presentan muchos aspectos problemáticos que son de dudosa validez y generalmente falsos. Estos enfoques aparecen en pares que se han dado en llamar "naturalista versus experimental," "psicológico versus sociológico," "cuantitativo versus cualitativo," y "el todo versus las partes."

Un aspecto realmente importante es constatar si un determinado enfoque está relacionado o no con la pregunta a la cual el investigador trata de buscar respuesta. Y lo que es más significativo aún, es si realmente vale la pena efectuar la pregunta.

El *rigor*, con un énfasis en la constatación de hechos por medio del uso de métodos exactos, es comparado con el *vigor*, que pone un énfasis en métodos que recogen el sabor de la vida en su totalidad. Esta dicotomía se usa para ilustrar el hecho de que la elección de un método apropiado depende de la naturaleza del problema que se va a investigar, de la etapa de desarrollo en que se encuentre la correspondiente área de investigación, el problema específico sobre el cual se hará la investigación y el estilo preferido del investigador.

El valor de la investigación dependerá directamente de cuán sistemático e imaginativo sea el enfoque. Aparte del pensamiento conceptual y una teoría

relacionada, o de una evidencia acumulada que lo respalde, ningún resultado de investigación en nuestro campo tiene sentido en forma aislada.

RESUMO

Na controvérsia entre "rigor" versus "vigor," muitos aspectos se confrontam. Estas discussões são geralmente de carácter duvidoso, tal como naturalista versus experimental, psicológico versus sociológico, quantitativo versus qualitativo e todo versus partes.

O problema real é se a abordagem específica que está sendo usada, é adequada para responder a pergunta proposta. Ainda mais importante é saber se a pergunta em si, merece consideração.

Rigor, ou a ênfase na constatação de fatos através de métodos exatos, é comparado com *vigor*, ou a ênfase em métodos que se propõe a capturar o sabor da vida na sua totalidade. Esta dicotomia é usada para ilustrar o fato de que a escolha de um método apropriado depende do problema a ser pesquisado, do nível de desenvolvimento do campo de pesquisa em si, e o estilo preferido pelo investigador.

O valor da pesquisa dependerá de quão sistemática e imaginativa for a abordagem, pois nenhuma evidência experimental em nosso campo poderá ser significativa em si mesmo, aparte do pensamento conceitual, existentes teorias e dados relacionados.