Toward An Epistemology for Applied Behavioral Science

Lee G. Cooper
Graduate School of Management
and
Harold G. Levine
Graduate School of Education

University of California, Los Angeles
Organzational and Behavioral Science
Working Paper No. 80-7
September 1980

ABSTRACT

Increasing numbers of social scientists carry out their work under the general rubric of applied behavioral science. All produce evidence which appears to bear upon the larger questions of human behavior. Until quite recently the connection of this evidence with that provided by "real" science has been necessarily problematic, in large part because of the philosophy of science which governed empirical research. In this paper we review both the Received View and Weltanschauung approaches to science, and show how critiques of these approaches require a new epistemology for scientific research. Beginning with the writings of Federick Suppe we develop the outlines of such an epistemology at the same time demonstrating that applied behavioral science ought to have the same foundation. The key points in this epistemology are the distinction between scientific belief and scientific truth and the crucial role of theory and domain development within an area of scientific research. We illustrate these principles with an extended example of the teaching profession as one kind of applied behavioral science.

Toward An Epistemology for Applied Behavioral Science

Introduction

Applied behavioral science covers many domains of activity: counseling and therapy of many kinds, organizational development and change efforts, ongoing management, program evaluation, consumer behavior, and education. While some workers in these domains contribute to the development of scientific knowledge, many others might express the view that the nature of their commitment to ongoing professional activity precludes them from involvement in science. Others may feel that "science" is not interested in what they are doing, or science is experimentation and they should not subject their clients to it. Although such views are understandable derivatives of the canons which influenced much of the thinking of philosophers of science until the end of the 1960's, a decade of great progress in philosophy of science has provided new opportunities for applied behavioral science.

Changes in the Philosophy of Science

In 1969, the University of Illinois convened a synposium to address the question," "What is the structure of a scientific theory?" The main proponents and critics of the traditional view of scientific theories and advocates of the major alternative views presented their positions and debated opposing views. The implications of these discussions for applied efforts became clear with the publication of the second edition of the proceedings (Suppe, 1977). This volume presents not only the position papers and a summary of the discussion, but also a critical introduction and afterword by Professor Suppe. The latter two sections are of book length and together document the traditional philosophy of science view, the critique of this view current at the time of the symposium, and the "rather dramatic developments in the philosophy of science since the 1969 symposium" (p. iii). Our very abbreviated remarks in the following section are based mainly on Suppe's writings.

The Received View

The Received View is the product of logical positivism -- the predominantly Germanic movement which sought to cleanse philosophy of abstract metaphysical speculation which did not allow for empirical specification. It was strongly influenced by Whitehead and Russell's Principia Mathematica (1910 - 1913), which attempted to reduce all of mathematics to logic; and, in its final form, was intended to be suitable to explicate all scientific theories. Although Whitehead and Russell's brilliant efforts fell short of that goal, the scientists, mathematicians and scientists-turned-philosophers of the Vienna Circle and Reichenbach's Berlin School were convinced by their arguments. The orthodox politics of the Germanic university system of the time made it less likely that alternative views would be developed. The result was the original version of the Received View which held:

> A scientific theory is to be axiomatized in mathematical logic... The terms of the logical axiomatization are to be divided into three sorts: (1) logical and mathematical terms; (2) theoretical terms; and (3) observational terms which are given a phenomenal or observational interpretation. The axioms of the theory are formulations of scientific laws, and specify relationships holding between the theoretical terms. Theoretical terms are merely abbreviations for phenomenal descriptions (that is, descriptions which involve only observational terms).
>
> (Suppe, 1977, p. 12)

A theoretical term was explicitly defined to be nothing more than what is observed of specified phenomena or their properties and correspondence rules were used to formalize the relationship of the two. This procedure enables proponents of the Received View the opportunity to purge scientific theory of metaphysical entities, since these entities were not considered to be either phenomenal or observational.

Once metaphysics had been eliminated from theory, philosophers began to ask whether it could not also be stricken from the language of science. The issue was how the terms of a scientific language would be related to the phenomenal world so as to be empirically verifiable. One could directly accept a person's description of phenomenal experience as an observational term or only allow a physicalistic language where one speaks of the observable properties of material things. At the time, standard doctrine held that a person's reports of sensory experience were error-free and thus presented no verification problem. Nonetheless physicalism won out, and the Received View restricted observational terms to material things and their observable properties.

A major modification of the Received View resulted from the recognition by Carnap (1936-37) that the explicit definitional form did not work for dispositional terms. Using Suppe's example of the dispositional term "fragile" we have:

> An object x is fragile if and only if it satisfies the following conditions: for any time t, if x is struck sharply at t, then x will break at t.
>
> (p. 18)

The truth of the phrase "an object is fragile" rests on the truth of the conditional phrase "for any time t, if x is struck sharply at t, then x will break at t." This conditional phrase is considered true whenever the antecedent (i.e. if x is struck sharply at t) is true <u>and</u> the consequent (i.e. x will break at t) is true. But the phrase is also considered to be true whenever the antecedent is false (c.f. Quine, 1953). Thus an object is considered fragile if it is not tested. This ridiculous state of affairs is true for <u>all</u> operational definitions, which prompts Suppe to comment in a footnote:

> It seems to be characteristic, but unfortunate, of science to continue holding philosophical positions long after they are discredited. Thus, for example, Skinner's radical behaviorism, which insists on operational definition came into prominence and dominated behavioral psychology well after most philosophers had abandoned the doctrine of operational or explicit definitions; taxonomists today strongly insist on operational definitions for taxa...
> (p. 19)

It should suffice that operational definitions are logically inadequate, but they are practically inadequate in addition. The dictum that a concept be synonymous with the set of operations used to measure it does not allow for measuring the same concept by different means. Different ways of measuring lead to different concepts according to operational definitions. Physical sciences could not progress without different ways of measuring the same thing, and neither could social sciences. Yet how many applied behavioral scientists still believe that if they wish to be "scientific" they must operationally define their terms?

The Received View replaced the explicit definition with the <u>reduction sentence</u> which partially defines a theoretical term by specifying sufficient but not necessary conditions for a particular instance to be considered an application of a theoretical term. Using the previous example this means that for the object x at time t, striking it sharply implies it will break if and only if it is fragile. This allows a nonfragile item which is not struck to still be nonfragile -- avoiding the logical flaw called the contrafactual conditional, while also allowing for other ways of measuring fragility. If the object were twisted sharply it would break if and only if it were fragile. All kinds of tests for fragility could be included by further reduction sentences, each adding partially to the meaning of the concept.

But even these progressive revisions of the Received View could not completely rescue it. Not all of the meaning of all theoretical terms can be introduced by reduction sentences consisting solely of observational terms or by terms completely defined by observational terms. On one hand theoretical terms could be considered to refer to real, nonobservable entities, processes or states. This realist interpretation of theoretical terms or hypothetical constructs (c.f. MacCorquodale and Meehl, 1948) implies that terms will have substantive meaning beyond that which can be encompassed by the Received View: intelligence implies more than that which is measured by a standard intelligence test; anxiety is more than the score on a manifest anxiety scale. On the other hand one could deny that theoretical terms refer to any nonobservables which really exist, and assert they refer to convenient fictions--intervening variables (MacCorquodale and Meehl, 1948), as in the instrumentalist interpretation--which can be dealt with by the Received View. The average family with its 2.1 children and 1.7 televisions exemplifies such a convenient fiction. The instrumentalist interpretation, however, leads to the theoretician's dilemma (Hempel, 1958), which holds that if theoretical terms serve their purpose, they are not necessary; and if they do not serve their purpose, they are surely unnecessary. Kendler's (1952) and Skinner's (1950) discussions of whether or not theories of learning are necessary are examples of this style of argumentation in psychology. While Hempel's dissolution of the theoretians shows that theoretical terms are necessary even for the instrumentalist interpretation, the instrumentalist is left in the uncomfortable position of needing theoretical terms while acknowledging that they do not mean anything (c.f. Suppe 1977, p. 34).

The meaning of theoretical terms is not wholly observational. "To fully specify the meaning of (theoretical) terms, recourse must be made to a richer metalanguage" (p. 35). Even if such a position were reconcilable with the Received View, the "coup de grace" is administered in a series of direct attacks by Putnam (1962) and Achinstein (1965, 1968) reported by Suppe (1977). They demonstrate that the features of direct observation which are intended, in the Received View, to differentiate observable from theoretical terms do not hold. Observation, Achinstein notes in Suppe (1977), has the following characteristics:

> (1) how many aspects of an item, and which ones, I must attend to before I can be said to observe it will depend upon my concerns and knowledge;
>
> (2) observing involves paying attention to various aspects and features of the item observed, but does not always require recognizing the kind of item being observed;
>
> (3) it is possible to observe something even though it is in a certain sense hidden from view -- for example, a forest ranger observes the fire even though he can only see smoke - so observing an item does not necessarily involve seeing or looking at it;
>
> (4) it is possible to observe something when seeing an intermediary image -- for example, when looking at myself in a mirror;
>
> (5) it is possible to describe what I am observing in the sky as a moving speck or an airplane.
>
> (p. 81)

The first and second points indicate that observation is preconditioned by cognition and intention. The third and fourth points deal with the role of implicit causal theories in observation: one infers fire from smoke and one assumes the image in the mirror is nondistorted. The fifth point emphasizes the conscious choices an observer has regarding description. All point toward the "theory ladeness" of observation (i.e. observational terms are tainted with theoretical terms as antecedents). These characteristics do not show that the observational-theoretical distinction is untenable but rather that

> ... the distinction has not been successfully drawn, and what is more, cannot be drawn in any plausible way on the basis of ordinary usage of terms in natural scientific languages. The only way the distinction could be drawn is artificially in a reconstructed language, and doing so would introduce an unwarranted degree of complexity into the analysis. Furthermore, even if the distinction is drawn satisfactorily it will mark no philosophically significant or epistemically revealing distinction. Finally, the distinction fails to capture what is distinctive either of theoretical terms or observation reports in science. The observational-theoretical distinction obviously is untenable. As such most of the epistemological interest of the Received View is lost. Insofar as the observational-theoretical distinction is essential to the Received View, the Received View is inadequate.
>
> (pp. 85-86)

Although a good deal of further criticism of the Received View is provided in Suppe and elsewhere, it suffices to say that the view which construes scientific theories as axiomatic calculi in which theoretical terms are set apart from observational terms and given meaning by operational definitions, is no longer taken seriously.

<u>Weltanschauungen</u> Approaches to Science

If science does not proceed as the Received View would hold, how does it proceed? Kuhn (1970) holds that science consists of periods of normal science interrupted by occasional scientific revolutions. Normal science "means research firmly based upon one or more past achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice" (1970, p. 10). The theory or collection of theories the scientific community shares is called a paradigm or disciplinary matrix. Normal science proceeds by elaborating the paradigm and solving the puzzles raised by it. Eventually a puzzle which is unsolvable within the paradigm arises and causes scientists to consider alternative theories. "The revolutionary proliferation of alternative theories continues until one of them emerges as victor, and a new scientific community coalesces around, and gives allegiance to, that theory -- at which time normal science reemerges" (Suppe, 1977, p. 636). Kuhn maintains that this is not only an accurate record of the history of science, but also how it ought to be. Feyerabend, according to Suppe, doubts that "normal science" really exists. He believes that science ought to progress via a proliferation of competing, incompatible theories.

Despite differences in what Kuhn and Feyerabend say "ought" to occur in science, their epistemologies are very similar and similarly problematic. To both of them scientific knowlege is based on nothing more than the sociological or psychological agreements among a group of people who call themselves scientists. Both positions lack a reasonable basis for evaluating what is progress in science. Without guidelines for evaluating progress, the replacement of one paradigm by another is merely a sociological event.

Philosophers of science have moved away from Weltanschauungen views for several reasons. Kuhn's notion of "paradigm" has always been vague. Masterson (1970) provides twenty different definitions of "paradigm" which all appear to fit Kuhn's view. Its replacement with the notion of "disciplinary matrix" has not clarified its meaning. The styles of definition, or correspondence rules he espoused became more and more like the Received View as Kuhn modified his position to respond to his critics. But how are we to evaluate progress in "normal science?" How are we to proceed in the absence of puzzles the paradigm cannot solve? Kuhn shortchanges the role of rationality in the growth of scientific knowledge. There is also growing skepticism that Kuhn's historical view of the cycling of science between normal and revolutionary periods is an accurate portrayal of history. But most basic to the criticisms is that Kuhn espouses an epistemology so subjective that "it makes discovering how the world really is irrelevant to scientific knowledge..." (Suppe, 1977, p. 648).

An Epistemology for Applied Behavioral Science

The "standard epistemological view" of knowledge is as a "justified true belief." That is, one knows a proposition is true if and only if:

(a) the proposition is true;

(b) one believes it is true; and

(c) one has adequate evidence for believing it is true.

In addition, Suppe (1977, p. 717) argues that it has been at least tacitly assumed in most recent epistemological writings that part (c) embodies what has come to be known as the "K-K thesis" Hintikka, (1962):

(d) one's knowing a proposition entails that one knows that one knows the proposition.

This thesis, in effect a kind of philosophical "reflexivity," implies that one cannot know the "truth value" one assigns to theoretical propositions about real-world phenomena unless one also knows the "truth-value" of the claims that are made as to the veracity of these same theoretical propositions. Thus, the evidence used to justify one's belief in the truth of one's claim to knowledge. Acceptance of the K-K thesis presents insurmountable obstacles to advancement in the observational sciences for two reasons. First, as we have already argued the meaning of theoretical terms need to be entirely observational in nature. As a result, it would be impossible to regularly specify all exact correspondence rules which link observational dimensions of the phenomenal world to theoretical terms. Second, the level of certainty is so demanding that even rigorous statistical tests fail to satisfy the criterion (see below).

If we accept Suppe's challenge (1977, p. 725) to develop an epistemology which eschews the "K-K thesis," we can identify three critically important benefits which have direct implications for an epistemology of social science, and, more importantly for our purposes here, for an applied behavioral science:

(1) Evidence gathered by scientists can be used either to justify knowledge or to help defend one's assertions that the body of knowledge in question is truly known. To use somewhat different terms we can distinguish between, on the one hand, what scientists believe to be true and, on the other, what evidence scientists may use to justify their claim that such-and-such is true. This distinction permits scientists to continue on in their work without having all the evidence one would need to guarantee one's claim to certainty of knowledge. Thus a phenomenon can be known under limited circumstances without the necessity of defending one's claim to know the phenomenon (Suppe, p. 722).

(2) The distinction made in (1) above seems to capture the essence of the actual process by which most scientists come to judge each others' claims to knowledge and by which they augment the corpus of scientific knowledge.

(3) Scientific evidence can never be totally confirming or disconfirming of a scientific theory in the sense that it represents a crucial test of some aspect of that theory. That evidence, rather, augments what scientists already believe to be the case about the phenomena in question, and only later, perhaps, becomes part of the evidence used by scientists to make claims about the scientific truths of their propositions about the phenomena.

The Bases of Scientific Beliefs

(1) Inferential statistics:

Inferential statistics have often been presented as if they provided enough evidence to assert that a proposition is true, but this is not true.

A basic test of a hypothesis is subject to two kinds of error. Type I error is the probability of rejecting a null hypothesis when one should not. The level of this type of error is routinely set at one out of twenty or one out of one hundred. While one may be willing to assert a belief in a proposition if one is wrong only one time out of one hundred, this is far too liberal a criterion on which to assert the truth of the proposition. Of course if this were the only serious threat to accepting a proof one could simply set the acceptable level of Type I error at some very extreme value, say one chance of error out of ten billion, and tolerate the residual uncertainty. Even in elementary statistics classes we were all taught about Type II error, the probability of failing to accept a correct alternative explanation (hypothesis). The Neyman-Pearson Lemma for uniformly most powerful tests establishes for us the standard procedures for testing a hypothesis so as to minimize the probability of Type II error for all possible alternative hypotheses. These procedures ensure, in general, that the more certain we insist on being, the less powerful are our statistical tests. So while inferential statistics can provide an adequate basis of scientific belief, it is not rational to use them directly as a basis of "scientific" truth.

(2) Observation:

The distinction between observational and theoretical terms, so prized by the Received View, has not been tenably drawn. However, to most reasonable people, seeing is believing. Shapere (c.f. Suppe, 1977, pp. 689-691) presents an analysis of observation in which he rejects the Received View notion on a theory-neutral observational language and avoids the relativism of observation explicit in the Weltanschauungen view. According to this analysis the starting propositions of these two positions are "plausible, and appear to constitute adequate criteria for any philosophy of science."

(i) Observation must be independent of, neutral with respect to, the theory to be assessed.

(i') Observation, if it is to be relevant, must be interpreted.

These basic principles are compatible with many approaches to applied behavioral science, including applied phenomenology. The willing suspension of disbelief prior to the onset of a phenomenological encounter of some sort, termed "bracketing," is an embodiment of the spirit of (i). The mulling over of events after an encounter with the phenomena, termed "sifting," is in the spirit of (i'). It leads to an interpretation which need not be preordained by theory.

Shapere demonstrates that the propositions which follow developmentally in the Received View and Weltanschauungen analyses do not follow logically, and thus we can accept these initial propositions without becoming locked into discredited positions.

We observe more than we can actually see. A traditional argument showing the limitations of observation comes from astrophysics where a whole theory of instrumentation (e.g. radio telescopy) interposes itself between the eye and the phenomena under "observation." Any reliable measurement provides at least as sound a basis for scientific belief as do statistical inference and direct observation.

The recent writing of Fiske (1979) on reliability deserves comment. He juxtaposes the study of characteristics of persons to the study of behaviors and discusses the implications of these studies in terms of reliability. Although he speaks of these studies as if they were worlds apart-- the former roughly corresponding to the building of response-response (R-R) laws and the latter roughly corresponding to the building of stimulus-response (S-R) laws-- applied behavioral science uses both kinds of studies. In particular contexts where the focus is on behaviors, the relatively short duration of actions, the recognizable onset and termination of behaviors, and the low level of abstraction of the terms referring to those behaviors, each acts to foster the reliability of observation and/or measurement. But the particularity of the context raises questions of ecological validity. That is, to what extent does that which is learned in applied behavioral settings (e.g. the clinic or the classroom) generalize to the daily lives of those involved.

The study of the characteristics of persons is portrayed by Fiske as focusing on entities, processes, or states of longer duration. A primary use of this kind of study in applied behavioral science entails the summarization of an individual's history as it applies to a current context. Stimulus-response approaches are ill-suited for such efforts. The relatively low interrater agreement cited by Fiske (1979, p. 35) does not typify the reliability of the broad class of methods for the development of response-reponse relations. Generalizability theory (c.f. Cronbach, Gleser, Nanda and Rajaratnam, 1972; Shavelson and Webb, in press), construct validity (c.f. Cronbach and Meehl, 1955), latent trait theory (c.f. Lord and Novick, 1968), and general multi-measurement approaches to causal modeling (c.f. Bentler, 1980) all entail entirely reasonable approaches to the establishment of scientific belief.

One restraint on the utility of information, if not the reliability, which deserves more attention is what is termed "method variance." In the study of characteristics of persons method variance refers to the component of information which results not direclty from the phenomena, but from how the phenomena are measured or assessed. In the study of behaviors method variance refers to the unique component of information contributed by a particular observer (e.g. the teacher, therapist or consultant). In general an applied behavioral scientist should attempt to keep track of the impact of his or her choices on the pheonomena under study.

The Role of Theory

Theory is the device by which scientists accumulate information resulting in belief in a domain. At different levels of abstraction theoretical statements allow for a parsimonious representation of the state of belief in a domain of inquiry. In Jungian psychology, for example, a high level theoretical statement might be that personality is manifested in two functions-- how an individual perceives and processes information about the world-- and in the attitude of whether one looks inside or outside oneself for such information. In a sensitivity training context a related theoretical statement might be that individuals whose thoughts dominate their feelings will gain more appreciation for others' feelings if they are simultaneously supported for the quality of their thoughts and confronted for their absence of feeling (c.f. Taylor, Note 1). In an organizational development context a high level theory might say that an organization must periodically regenerate its reason for being or cease to exist (Mittler, Note 2). This theory results in a set of diagnostic signs for three cyclic states of organization life.

Such statements in a domain are part of a relational network. In the network there will also be some reduction sentences, which partially define hypothetical constructs in terms of reliable information; statements of direct observation; statements relating the results of research efforts to theoretical constructions; and statements which relate other theories to ones currently under consideration. For convenience we can consider that some sentences, phrases or terms are hypothetical and others are evidential. The network in a domain consists of four kinds of relations: hypothetical to hypothetical, evidential to hypothetical, hypothetical to evidential, and evidential to evidential.

Shapere (1977) recognizes three kinds of theoretical inadequacies. The first is incompleteness. If the network of relations which covers the domain is thought of as a mesh or sieve, incompleteness refers to a hole or break in the mesh -- an area of the domain which is not covered as completely as others. A theory of orienting style (Taylor, Note 1) may give a very useful explanation of the temporal changes brought about by differing normative profiles of group members, but it provides no account of possible permanent changes in personality structure. The second is simplification. A sieve is a three-dimensional object, the phenomena one wishes to cover with a network may well be four-dimensional or more. As an individual changes in response to the normative profile of a group, the normative profile of the group also changes. The theory is simplified by not attending to the change in group profile over brief periods of time. An organization is a multi-group, multi-facet structure. Different parts of the organization may be in different substages of the regenerative cycle at a given point in time. Any major division will still simplify the process occurring at sublevels to a certain extent. The third kind of inadequacy he calls black box incompleteness. This refers to the fineness of the mesh or sieve. No matter how fine, the mesh is never closed and something may always pass through. There is system and structure to the controlling influence of each individual's unconscious on his or her orienting style. At the present state of development of the domain this is largely an incomplete "black box."

The Basis of Scientific Knowledge

The past reliance on operational definitions can be viewed as being based on an unworkable premise that all scientific knowledge is stipulative. Our overconfidence in hypothesis testing can be viewed as being based on an unworkable premise that all scientific knowledge is inferential. If one is to find a workable premise for a thing as broad as "all scientific knowledge" it had better be very general.

The most general premise is that all scientific knowledge is relational. This conforms to our everyday experience since belief and knowledge -- scientific or not -- are relations of one sort or another. Most of what we know about worldly objects comes from our history of relations with them. Most of what we know about things we cannot see comes from how they relate to things we can. Inference is a particular relation we use sometimes as a basis for knowledge. It is not that such a relation is disallowed as a basis for scientific knowledge. Rather it is the primary or exclusive use of it which will lead to difficulties in science.

Consider the possiblity that knowledge and truth do not accrue to a particular relation in isolation. A relation is a statement. Out of context it is rarely, if ever, possible to assert the truth of a statement. What then is the context of a stated relation which allows us to speak of knowledge and truth? Up to this point we have used the term "domain" in the sense in which we are all _a priori_ familiar. But Shapere's formal use of "domain" provide the context which allows us to bridge from belief to knowledge (c.f. Suppe, 1977, pp. 521-557 and 686-704).

Items of information become associated together as bodies of information which have the following characteristics:

(1) The association is based on some relationship between the items.
(2) There is something problematic about the body so related.
(3) That problem is an important one.
(4) Science is "ready" to deal with the problem (p. 525)

Such bodies of information Shapere calls domains. That the boundary of a domain may be fuzzy or that the division and coalescence of domains into subdomains and superdomains may not be clearly regulated, has no detrimental impact on our discussion.

The items in a domain need not be classified as theoretical or observational. The four kinds of relations mentioned in the section on the role of theory are adequate. Relations of these sorts form networks in a domain. These networks replace the need for the observational-theoretical distinction.

The relational network in a domain carries the largest burden as a basis for scientific knowledge. Evidence from all sources can increase or decrease our belief in the propriety of an asserted relation. Confirming or disconfirming evidence have equal status in terms of their impact on belief in the network. Both are partial, weak, fallible, local and temporal. However evidence does flow in a network from areas of higher concentration to areas of lower concentration. If we know a method is proper, this increases our belief in unusual or unexpected results in some application. If we know what results are proper, our belief in new methods increases if they produce these results.

A relational network may stem from a simple statement of high level theory. While high level theory may be valued for its elegance and simplicity, the relational network is valued for its complexity and intricacy. The abundance of confirmed relations between hypotheticals and evidentials is what allows us to overcome the inherent weakness of a single experimental/research result in isolation.

Competitive support and parsimony also have roles in the establishment of scientific knowledge. It is rare, yet possible, that competing theories are specified in ways such that all evidence relating to one theory can be evaluated against all the evidence relating to the other theory. If this occurs in richly articulated domains the theory offering the better total explanation is accepted. For explanations with equal support the more parsimonious one is accepted. But competitive support and parsimony are often over sold. Garner, Hake and Ericksen (1956) build a system called "convergent operationism" from the sense that competing theories could be evaluated by a series of "crucial experiments" which would result in one final contest. This final experiment would supposedly crucially eliminate all but one theory. But, in reality, there will still be an infinite number of alternative explanations for the result in the final experiment. Whole networks may at times be compared to other networks, but sequential elimination by successive experimentation is untenable. In young domains, such as the behavioral sciences, parsimony is easily misused. It is thus that theories of memory without organization or choice without values might be advanced in the name of parsimony prior to having the appropriately specified alternative theories for competitive evaluation.

Theories in a domain become scientific knowledge when the overwhelming burden of evidence supports the hypothetical relations. This does not mean there must be no disconfirming evidence. If this were so we would attribute more value to such evidence than it deserves. Once the entire network has reached an extreme level of confirmation, the theories in the network are considered to be lawful relations which may be used as evidence in support of hypothetical relations in other domains if a basis for the generalizability of the theories is established. One may judge the completeness of a domain by the resolution of the problems which characterize it. The professional who is aware of the completeness of a domain may rationally judge when there is sufficient confidence in the asserted relations to use them as a basis for experimental practice, when there is enough confidence to establish policy and when there is enough evidence to assert theoretical relations as knowledge.

The Role of the Practitioner: Teachers as Scientists

Though practitioners may not normally regard themselves as "researchers" there are several reasons why their efforts are ideally suited to contribute to science. In the first place, the standards of professional responsibility which guide the practitioner are consonant with the best principles of scientific inquiry: learning as much as possible about the phenomenon in question and generalizing this knowledge to broader sets of circumstances; maintinaing well-kept records of one's efforts and results; and adopting a stance which values "listening to" rather than "listening for" and "looking at" rather than "looking for." Second, applied behavioral scientists, as groups of people working in a particular subject area, have access to a far greater population of individuals and situations than any "researcher" is likely to obtain and, therefore, are more likely to be able to observe the full range and diversity of the behavior in question. The various professional societies and publications of the applied behavioral scientists are the fundamental mechanisms for the support and dissemination of these efforts.

An implicit consequence of recent thinking in philosophy of science is that the epistemological base for conceptualizing "traditional" social science is no different than that which underlies applied behavioral science.

In fact, philosophy of science is currently modeling the very process which that discipline claims as its subject matter: the justification of scientific practice and knowledge. Science is increasingly coming to be practiced by the applied behavioral scientists. Though such scientists typically phrase their research questions in terms of what works, and why, and are often eager to try to make something work, they do come upon evidence pertaining to one or more domains of knowledge. The more traditional scientist confronts this same or related evidence though his or her research question is most often concerned with what is happening and why. Under the Received View, with its seemingly inviolate distinction between observational and theoretical terms, its operational definitions and extensive controls, its "covering-law" explanations and axiomatized theoretical propostions, and its crucial tests and universalist "proofs," efforts of applied behavioral scientists were often outside the process by which knowledge was considered "justified." However, post-positivistic science recognizes domains of scientific interest where both theory and observational fact are not always distinguishable and are, in any case, mutually interdependent, and where the role of theory is as crucial to scientific belief as it is to scientific truth. The well-trained applied behavioral sceintist, as much as any "pure" researcher, is both a user of and a contributor to theory and hence to science. We wish to illustrate more explicitly how this might be done by using the example of teachers of retarded learners.

Any practicing teacher is a professional whose work combines both observational and communicational skills in the classroom with the best of current instructional and learning theory. Some of the main points of the theory which informs the efforts of both teachers and academicians toward retarded learners[1] are listed below:

(1) Learning is directly related to mental age. The general laws of learning which apply to "normal" children apply to retarded children as well.

(2) While initially haphazard, performance once the child understands what to do, tends to occur at the same rate as normal students of similar mental age.

(3) Thought processes of retarded students are described as concrete, discrete, unrelated, immediate, and obvious.

(4) Retarded learners have poor discrimination abilities so that similar materials and competing stimuli impede learning.

(5) Transfer of training is extremely limited and may only regularly occur when transfer is by identical elements rather than by principle.

(6) Self-teaching or auto-instruction is typically deficient.

(7) Incidental learning seems deficient though the role of cultural and experiential factors is unknown but is suspected.

(8) Learning strategies are more likely to be characterized by avoiding failure rather than by seeking success.

[1]We use the definition of mentally retarded learners given by the American Association on Mental Deficiency: individuals who possess significantly subaverage general intellectual functioning as well as deficits in adaptive behavior, both of which are manifested during the development phase of the individual's life (AAMD, 1973, p. 5).

(9) Retarded learners are also socially inept which has a further effect on their learning by narrowing the opportunities for learning or by confusing the learners due to misconstruals of the task demands (Ellis, 1979; Gearheart and Litton, 1975; Kolstoe, 1972, 1976; MacMillan, 1977).

These statements clearly are not a set of logically-derived or formulated propositions. Rather most express relations between hypothetical and evidential entities, process or states. As these items clearly relate to one another, are of importance and current concern among researchers, and are problematic in that all, in some way, attempt to explain what distinguishes retarded learners from normal learners while attempting, sometimes implicitly, to account for these differences, they seem to conform to Shapere's (1977) definition of a domain in scientific inquiry. Since "proof" for any of the elements in this domain is lacking--indeed, seemingly contradictory information is available for most--they comprise the corpus of scientific beliefs for most of those who work with retarded learners.

How might classroom teachers contribute to the elucidation of this domain? Since we have argued that the efforts of teachers and other applied scientists have no different epistemological status than "pure" research the question can be rephrased in more general terms: "What are the scientific problems connected with a domain and with the theories which help constitute that domain?" In general there are three sorts of scientific problems: domain problems, which involve clarification of the elements within and the boundaries of the domain; theoretical problems, which concern the level and type of explanation of phenomena within the domain; and problems of theoretical inadequacy which refer to the thoroughness of the theory itself.

Teachers of retarded learners, in their "new" role as applied behavioral scientists, are in an ideal position to contribute to our knowledge about this domain.

(1) Domain problems. In large part the history of research in mental retardation parallels that in psychology in general (Brooks and Baumeister, 1977). Consequently one of the major topics in mental retardation research is learning, particularly as studied behaviorally. Self-monitoring, incidental learning, generalization, discrimination abilities, and other elements of this domain have been characterized behaviorally and most often studied under experimental or quasi-experimental conditions. Teachers may now add a naturalistic perspective to the study of these elements and refine or reconceputalize them based upon their observations. Perhaps some expansion of the domain will be necessary to include meta-level cognitive abilities (Brown,1975; Flavel, 1971; Flavell and Wellman, 1977) or other phenomena as well.

(2) Theoretical problems. Two kinds of explanatory theories in science are compositional and evolutionary (Shapere, 1977, pp. 534ff.). A compositional theory explains by recourse to constituent parts of the individual units studied within the boundaries of the domain and regularities which govern these separate parts. Evolutionary theory, on the other hand, resorts to the development of individual units within the domain. Deciding on which perspective best fits the domain items in question, if indeed one can be said to be a best fit at all, is itself an important problem. Certainly it is a problem which is implicit in studying learning abilities of mentally retarded students. Teachers who often have the same students over a period of several years and who, in any case, maintain files and records of the children's progress (especially now that

individualized educational programs and evaluation are mandated by law in some states) are perhaps best able to make judgments and collect data relative to a "developmentalist" versus "atomist" explanatory framework.

(3) Shapere (1977, pp. 557-565) discusses three types of theoretical inadequacies: incompleteness, simplification, and black-box incompleteness.

(a) Incompleteness. A theory can be considered complete or incomplete relative to the body of information within a domain or, more usually, within a subdomain. In the area of mental retardation research recent findings indicate that the laboratory based tasks by which we typically evaluate such cognitive abilities as memory may be quite different in terms of their underlying structure than memorial task found under "everyday" conditions (Levine, Zetlin, and Langness, 1980). In addition, long-term observations of retarded learners reveal quite sophisticated social skills, high and enduring motivation toward specific goals as well as complex, shifting strategy selection in pursuit of these goals, widely varying performance incentives, and common use of symbols (e.g., abstract thought) and fundamental deductive logic in problem-solving. Current theory does not adequately provide for these manfiestations of cognitive functioning.

(b) Simplificaton. We can identify two different types of simplification. The first results from making assumptions about the units in the domain or their behavior when evidence is lacking, contradictory, or inadequate. The debate over whether trainable mentally retarded learners (those whose IQ fall in the 36-51 range) can profit from academic subjects is one such issue. There are some who maintain that trainable mentally retarded learners will always remain trainable, not educable, and therefore should be taught with non-academic curricula (c.f., Burton, 1974). Others maintain just the opposite and cite research to show that complex cognitive skills such as reading can be taught (c.f., Brown and Perlmutter, 1971).

The second form of simplification results from having general agreement among most researchers that they do possess the adequate evidence or theory (usually as part of another domain or theory rather than the one of immediate concern) for dismissing certain "background information" as irrelevant to the issue at hand. In fact, this claim is an hypothesis which may prove to be quite erroneous as further research accumulates. The history of classroom research shows increasing recognition of the diversity and multiplicity of variables which affect learning outcomes: contextual variable affecting tasks and task presentation, indivdidual variables such as motivation and the student's perception of the demands made upon them, teacher expectations, and the like.

(c) Black-box incompleteness--In (a.) above the incompleteness is a "known" quantity--i.e., the researcher knows what has been omitted and realizes its crucial role in the theory or domain. Black-box incompleteness refers to a state in which the attributes of some more micro level phenomenon may play a role in explanation, but just how and even whether it does so remains unclear. In mental retardation research this problem is a major one. For many individuals labeled as mentally retarded we are unable to pinpoint its underlying cause or, even when this is possible, we cannot usually characterize the specific neurophysiological (or other) effects on the individual's cognitive abilities. The result is a true "black box" which is often glossed over, disclaimers to the contrary, by speaking of the trainable mentally retarded child as if there were some homogeneous group of such children. In addition it is unclear whether any lack of academic progress for the individual student is solely related to mental inadequacy or is equally traceable to other problems such as lowered teacher expectancies and/or lowered student motivation, inappropriate teaching strategies or instructional programs, or inappropriate social behaviors in the classroom by teacher and student alike which are not conducive to learning (e.g., teachers coddling their students engaging in infantile behaviors which, in turn, are tolerated by teachers).

Thus there are a number of ways in which teachers as applied behavioral scientists will make contributions to domains of scientific interest. Methodologically these contributions will come about through the classroom application of such research techniques as observational and diagnostic accounts of students focusing on discription, comparison, and explanation, simple frequency counts of characteristic behaviors, "natural" experiments (Sechrest, 1970), unobtrusive measures (Webb, Campbell, Schwartz, and Sechrest, 1966), accounts of unique or otherwise puzzling phenomena, and the like. Increasingly, N = 1 experiments (Dukes 1972; Hersen and Barlow, 1976) and interventions will become the common currency of action oriented research. Teachers can be expected to specify and measure the outcomes of new curriculum materials of teaching strategies as they try to make education work for their students. Results of this intervention must be compared with other such attempts by the same teacher, other teachers in the school, or other similar efforts reported in the literature. The teacher-researcher will look for commonalities, whether they be related to the contexts in which the techniques were employed, to the students themselves, to the kinds of tasks for which the new teaching method seemed particularly suited, or to the outcomes. Finding the commonalities is only the first step. They must be related to the statements within a domain and ultimately they must address one or more of the problems inherent to or subsumed within that domain. Their research, like that of "pure" behavioral scientists will neither prove nor disprove a theory. Rather, perseverance in their work with the problems of retarded learners will inevitably lead to evidence which (a) increases belief in an asserted set of relations, or (b) decreases belief in those relations, or (c) demonstrates the inadequacy of current theorizing and provides the basic datum for a new set of relations.

Conclusions

The question remains as to what guarantees the rationality in the growth of scientific knowledge. As it is with all human endeavors, science can not absolutely guarantee rationality. There are, however, very strong forces pressing for rationality. Most fundamental of these forces is the rationality of our purposes as deliverers of human services. The intention to solve a problem, to make something work, is not a liability. It is the driving mechanism for progress in science. As long as our intentions are to increase our understanding of and improve the human condition, the scientific knowledge which is a coproduct of those endeavors will grow rationally.

REFERENCE NOTES

1. Taylor, V. O. "Change of Orienting Style During Laboratory Training: A Jungian Perspective." Unpublished doctoral dissertation. Graduate School of Management, University of California, Los Angeles, 1977.

2. Mittler, E. "Organizational Regeneration." Unpublished doctoral dissertation. Graduate School of Management, University of California, Los Angeles, 1974.

REFERENCES

Achinstein, P. The problem of theoretical terms. American Philosophical Quarterly, 1965, 193-203.

Achinstein, P. Concepts in Science. Baltimore: Johns Hopkins Press, 1968.

American Association on Mental Deficiency. Manual on terminology and classification in mental retardation (Revised edition). Washington, D.C.: Author, 1973.

Bentler, P. M. Multivariate analysis with latent variables: Causal modeling. Annual Review of Psychology, 1980, 31, 419-456.

Bergmann, G. The logic of psychological concepts. Philosophy of Science, 1951, 18, 93-110.

Brooks, P.H. and Baumeister, A.A. A plea for consideration of ecological validity in the experimental psychology of mental retardation. American Journal of Mental Deficiency, 1977, 81, 407-416.

Brown, A.L. The development of memory: Knowing, knowing about knowing, and knowing how to know. In H.W. Reese (Ed.), Advances in Child Development and Behavior 10. New York: Academic Press, 1975, 103-152.

Brown, L. And Perlmutter, L. Teaching functional reading to trainable level retarded students. Education and Training of the Mental Retarded, 1971, 6, 74-84.

Burton, T.A. Education for trainables: An impossible dream? Mental Retardation, 1974, 12, 45-46.

Carnap, R. Testability and meaning. Philosophy of Science, 1936-37, 3, 420-468; 4, 1-40.

Cronbach, L. J., G. C., Handa, H. and Rajaratnam, N. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: John Wiley and Sons, 1972.

Cronbach, L. J. and Meehl, P. E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 218-302.

Dukes, W.F. N=1. In R.E. Kirk (Ed.), Statistical issues: A reader for behavioral science. Monterey, Ca.: Brooks Cole: 1972, 217-223.

Ellis, N.R. Handbook of mental deficiency, psychological theory and research. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

Fiske, D. Two worlds of psychological phenomena. American Psychologist, 1979, 34, 733-739.

Flavell, J.H. and Wellman, H.M. Metamemory. In R. V. Kail, Jr. and J.W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977, 3-33.

Garner, W. R., Hake, H. W. and Ericksen, C. W. Operationism and the concept of perception. Psychological Review, 1956, 63, 149-159.

Gearheart, B.R. and Litton, F.W. The trainable retarded: A foundations approach. St. Louis: C.V. Mosby, 1975.

Hempel, C. Theoretician's Dilemma. In Fiegl, H., Scrivens, M. and Maxwell, G. (Eds.) Minnesota Studies in the Philosophy of Science, Vol. 2. Minneapolis: University of Minnesota Press, 1958.

Hersen, M. and Barlow, D.H. Single-case experimental designs: Strategy for studying behavior change. New York: Pergamon, 1976.

Hintikka, J. Knowledge and Belief. Ithaca: Cornel University Press, 1962.

Kendler, H. H., "What is learned? -- A theoretical blind alley. Psychological Review, 1952, 59, 269-277.

Kolstoe, O.P. Mental retardation: An educational viewpoint. New York: Holt, Rinehart and Winston, 1972.

Kolstoe, O.P. Teaching educable mentally retarded children. New York: Holt, Rinehart and Winston, 1976.

Kuhn, T. S. The Structure of Scientific Revolutions, Second Edition, Enlarged. Chicago: University of Chicago Press, 1970.

Lekatos, I. andMusgrave (Eds.) Criticism and the Growth of Knowledge. Cambridge: Cambridge University Press, 1970.

Levine, H.G., Zetlin, A.G., and Langness, L.L. Everyday memory tasks in classrooms for TMR learners. The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 1980, 2, 1-6.

Lord, F. M. and Novick, M. R. Statistical Theories of Mental Test Scores. Reading: Addison-Wesley, 1968.

MacCorquodale, K. and Meehl, P. E. On the distinction between hypothetical constructs and intervening variables. Psychological Review, 1948, 55, 95-107.

MacMillan, D.L. Mental retardation in school and society. Boston: Little, Brown, 1977.

Masterson, M. The nature of a paradigm, pp. 59-90 in Lakatos and Musgrave (1970).

Putnam, H. What theories are not. In Nagel, E., Suppes P. and Tarski, A. (Eds.) Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress. Stanford: Stanford University Press, 1962, 240-256.

Quine, W. V. O. Methods of Logic. New York: Holt, 1953.

Sechrest, L. Experiments in the field. In R. Naroll and R. Cohen (Eds.), A handbook of method in cultural anthropology. New York: Columbia, 1970, 196-209.

Shapere, D. Scientific theories and their domains. In Suppe, F. (Ed.) The Structure of Scientific Theories, Second Edition. Urbana: University of Illinois Press, 1977, 518-565.

Shavelson, R. J. and Webb, N. O. Generalizability theory: 1973-1980. British Journal of Mathematical and Statistical Psychology, in press.

Skinner, B. F. Are theories of learning necessary? Psychological Review, 1950, 57, 193-216.

Suppe, F. The Structure of Scientific Theories. Edited with a critical introduction and afterword by Frederick Suppe. Urbana: University of Illinois Press, 1977.

Webb, E.J., Campbell, D.T., Schwartz, R.D. and Sechrest, L. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand McNally, 1966.

Whitehead, A. and Russel, B. Principia Mathematica, 3 Vols. Cambridge: Cambridge University Press, 1910-13.