1

History and Evolution

All things prepare the event. Watch.

— Eliot (1971, p. 183)

After reading this chapter, you should be able to define the following terms and provide the information requested in the study guide below.

TERMS

Human behavior
Science
Determinism
Conditioned response
Unconditioned response
Respondent conditioning
Operant conditioning
Behavior modification
Applied behavior analysis
Explanation (cause versus probability)

STUDY GUIDE

1. Provide a historical summary of the scientific activity that led to what we now know as the applied analysis of behavior.
2. Summarize the philosophical perspectives that provide foundations for scientific activity.
3. List and describe some of the main features of qualitative and quantitative research methods.
4. List some of the arguments for multiple research method compatibility.

5. Describe some of the myths about behavior analysis that have been popularized.
6. List and describe some of the reasons for the nonuse of behavior analysis data by applied professionals.
7. Explain what is meant by the fallacy of affirming the consequent when using large-group, statistically based research comparisons.

The experimental, and in particular the applied, analysis of behavior has a long and productive history. This is the case even though the method has suffered criticisms and related nonuse by many mainstream professionals in the education, psychological, and social sciences. Many experimental methods now appear in the applied literature, ranging from large group comparisons using a variety of statistical analyses to qualitative methods that focus on the use of existential explanations and subjective analyses. The main purpose of this text is to present a variety of behavior analysis and direct observational methods in the hope of arming faculty and students with important contributors to research activity. To give the reader a better understanding of behavior analysis as a research methodology, this chapter provides a brief history of behavior analysis as a research method, argues for its compatibility with other research methodologies, and debunks some prevalent myths that are detrimental to behavior analysis use.

In the education, social, and psychological sciences, behavior analysis has proved an important methodological contribution when research focus is on one or only a few study participants, and when generality is sought through careful and systematic replication of treatment implementation over time. The method has appeal over traditional large-group designs when another method to control for internal and external validity concerns is warranted. In relation to qualitative research methods in which a small group of individuals or a single setting is studied, behavior analysis methods have a thoroughgoing quantitative component to measure and evaluate treatment effects, and they include a rigorous set of data inspection procedures that are elegant in their simplicity and consistent in their application. Behavior analysis designs are also methodologically compatible with other research perspectives, lending themselves to enhancing the types of information gathered in a particular study. In many respects, then, behavior analysis designs provide an

important complement to other methods for conducting research and evaluation.

A BRIEF HISTORY OF BEHAVIOR ANALYSIS

Behavior analysis research designs have long been used in a wide variety of applied education, social science, and psychological disciplines and have produced a substantial scientific literature in many areas. The method was not formally recognized by name until the appearance of work by B. F. Skinner (1938, 1953, 1956) and the related creation of the Association for Behavior Analysis (ABA), and it has been characterized by terms such as *single-subject research, single-case methods, intrasubject replication designs,* and others (Kazdin, 1982). The method in various forms has been used in the applied sciences for over a century. Before we proceed to a more cookbook approach to implementing a behavior analytic study, we discuss where this methodology came from and what it is based on. In providing this sort of information, we hope that readers will develop a greater understanding of the scientific enterprise in general and be receptive to the use of behavior analysis when appropriate to their research questions of interest. Many contemporary mainstream scholars argue that behavior analysis is a radical departure from other research methods, but we hope to demonstrate that it is far more similar to other methods than it is different from them.

The brief history provided here is based largely on two sources (Kazdin, 1982, Chapter 1; Michael, 1991) that have encapsulated in summary form a chronology of the scientific methods contributions on which behavior analysis is founded. We have reorganized the information from these sources, placed our unique emphasis on various parts, and added historical and contemporary innovations by some often neglected contributors who should be included within a chronology.

The period from the late 19th century to the turn of the 20th century saw the beginning of single-subject and small-group research, and of what we now call the *applied analysis of behavior*. The work of Charles Darwin, specifically his foundational text *Origin of Species* (1859), is considered the first important origin of the methodology. Darwin's influence established the idea of the continuity of species. In other words, through a theory of human evolution by natural selection Darwin demonstrated that the study of nonhuman and human behavior is relevant to a greater understanding of a variety of cognitive,

perceptual, and evolutionary processes. The following quotation from Skinner (1984, cited in Michael, 1991) provides a summary of what Darwin posited:

> Human behavior is the joint product of (i) contingencies of survival responsible for natural selection, and (ii) contingencies of reinforcement responsibilities for the repertoires of individuals, including (iii) the special contingencies maintained by an evolved social environment. Selection by consequences is a causal mode found only in living things, or in machines made by living things. It was first recognized in natural selection: Reproduction, a first consequence, led to the evolution of cells, organs, and organisms reproducing themselves under increasingly diverse conditions. The behavior functioned well, however, only under conditions similar to those under which it was selected.
>
> Reproduction under a wider range of consequences became possible with the evolution of processes through which organisms acquired behavior appropriate to novel environments. One of these, operant conditioning, is the second kind of selection by consequences: New responses could be strengthened by events which followed them. When the selecting consequences are the same, operant conditioning and natural selection work together redundantly. But because a species which quickly acquires behavior appropriate to an environment has less need for an innate repertoire, operant conditioning could replace as well as supplement the natural selection of behavior.
>
> Social behavior is within easy range of natural selection, because other members are one of the most stable features of the environment of a species. The human species presumably became more social when its vocal musculature came under operant control. Verbal behavior greatly increased the importance of a third kind of selection by consequences, the evolution of social environments or cultures. The effect on the group, and not the reinforcing consequences for individual members, is responsible for the evolution of culture. (p. 477)

Essentially, Darwin subscribed to *logical positivism*, which is closely related to *scientific empiricism*. This school of thought limits propositions

either to those that are empirically verifiable through systematic experimentation or to those that are analyses of definitions and relations among terms. Empirically verifiable propositions became the concern of the scientific method, and analysis of definitions and relations between terms the specific task of philosophy. In addition, and promoted by 19th-century philosophers such as Auguste Comte, conceptual and theoretical focus is on the objective examination of all phenomena, a clear antimetaphysical and anticognitive approach to the explanation of events, and the belief in universal and primarily static laws that govern all events (Titus, Smith, & Nolan, 1986).

The foundations for a formal empiricism, or mechanical and experimental science, were also laid through Darwin's contributions. Within this foundation, the following defining elements were added over the course of the early 20th century by a variety of contributors. The following definitions are from Malott and Whaley (1983):

> *Science.* The systematic study of how events or occurrences are related to the production of other events or occurrences (p. 4)
>
> *Determinism.* The notion that an event is produced or results from the active presence of a precise set of conditions (p. 4)
>
> *Analysis.* The task of finding out and describing the precise set of conditions that give rise to specific events (p. 4)
>
> *Application.* The bringing together of the needed conditions in order to produce certain desired events or outcomes (p. 4)
>
> *Unconditioned response.* Response evoked by a stimulus even without prior experience (p. 109)
>
> *Conditioned response.* Response evoked by a stimulus only after pairing that stimulus with one that already causes the response (p. 109)
>
> *Respondent conditioning.* The procedure of pairing two stimuli with the result that one stimulus (conditioned stimulus) acquires the power to cause the response already caused by the other stimulus (unconditioned stimulus) (p. 109)
>
> *Operant conditioning*. A change in the likelihood of a response due to the results of that response (p. 93)
>
> *Learning.* A change in behavior that occurs as a result of experience (p. 93)

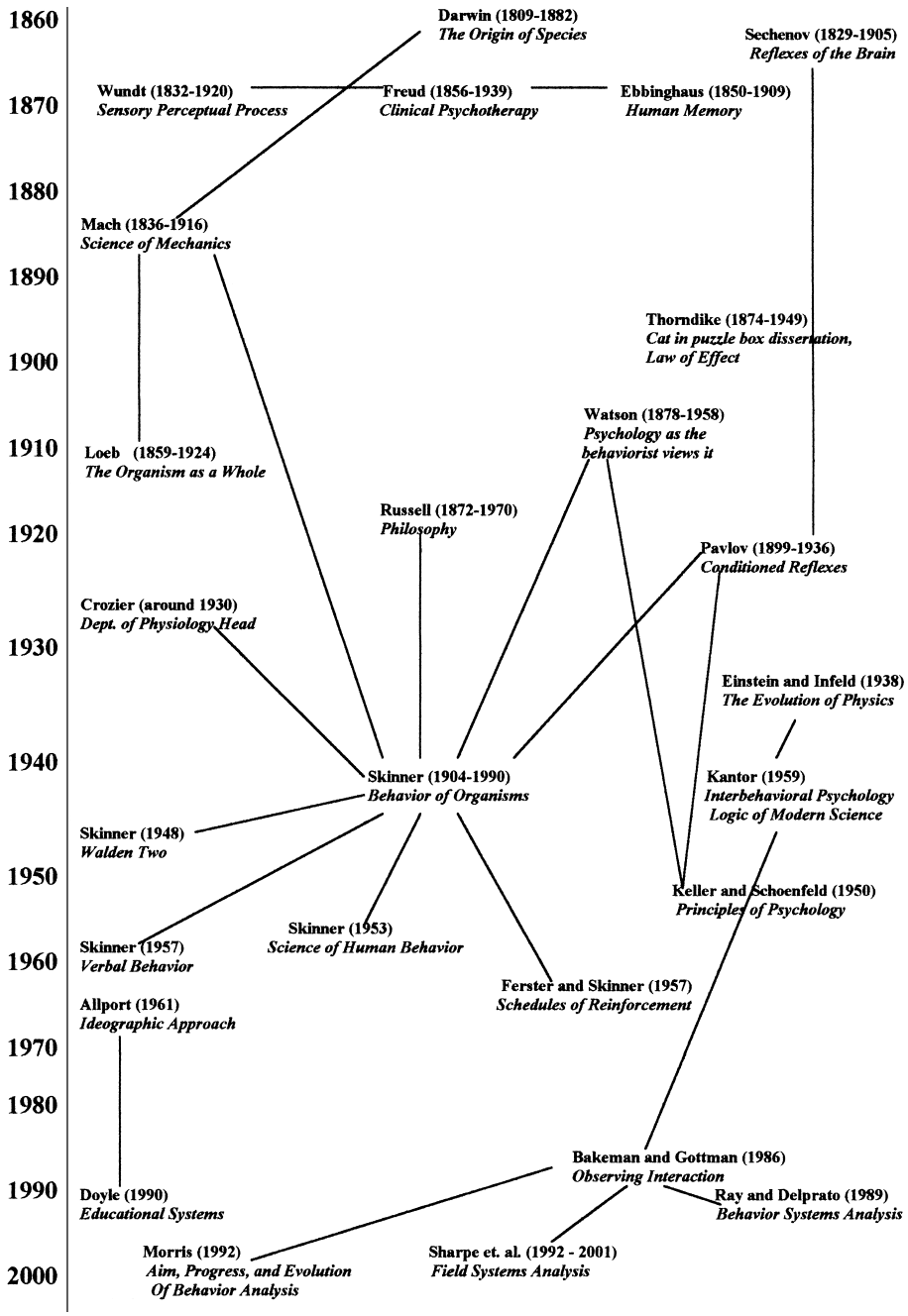
> *Behavior modification.* The application of principles of behavioral psychology, and operant psychology in particular, to change actions in a desired manner or direction (pp. 629-630)

Some of the important historical contributors to behavior analysis methods are highlighted in Figure 1.1. Although historical accounts are presented with variable emphasis and inclusion, we hope our account is useful (with potential instructor modifications).

A variety of scholars prominent in the psychological sciences at the turn of the 20th century were engaged primarily in research involving single subjects or small groups and provided important illustrations of the concepts included in the list of definitions above. Wilhelm Wundt, for example, engaged primarily in the study of sensory perception by reporting in qualitative and thoughtful ways the reactions of individual subjects to specific stimulus conditions presented to them. Hermann Ebbinghaus, a contemporary of Wundt, studied memory through analyses of individuals' learning and recall of nonsense syllables by altering stimulus conditions with individual subjects. Sigmund Freud, often considered the father of clinical psychotherapy by the popular culture, promoted early versions of in-depth study of individual cases designed to better understand and conceptualize basic psychological processes, developmental stages, functional relations among symptoms and historical characteristics of the patient, and a variety of cognitive processes he considered to be related to personality and behavior. A primary disadvantage of this type of early scientific activity, however, lay in its lack of experimental control, including unclear descriptions of independent variables or treatments given to subjects, oftentimes vague and poorly defined behavior or cognitive measurement systems, and limited documentation of particular subject characteristics. These types of problems often led to the limited or inappropriate generalization of findings to other subjects and population groups.

An important early behavioral contribution was that of Ivan Sechenov in his text *Reflexes of the Brain.* During a time period when most medical and physiological literature focused on the mental aspects of human thought and consciousness, Sechenov proposed that all aspects of cognition in humans were based on behavioral reflexes. He worked out a carefully designed explanation of complex human behavior through derivations of simple reflexes, their consequences, and their potential combinations that was very similar to more contemporary definitions of respondent conditioning. Through the theoretical model that Sechenov provided, a foundation was offered for a completely behavioral approach to the explanation of phenomena. This foundation

Figure 1.1 History of the Current Field of Behavior Analysis



was heavily relied on by the scientists who followed Sechenov, most notably Pavlov and a group of Russian physiologists.

Ernst Mach, a contemporary of Sechenov, provided through his work, *The Science of Mechanics*, a foundation for what we speak of today as a lineal mechanic approach to empirical research and the scientific enterprise. Mach first proposed that the nature and origins of scientific inquiry should be a natural outgrowth of the important practical concerns of daily life and that behavior is one important part of epistemological (i.e., how we know what we know) development. He essentially defined *science* as primarily behavioral, in that it is the objective and systematic pursuit of helping humans become more effective and productive. Mach next defined *cause* as a functional relationship among independent variables (treatments) and dependent variables (measures), and he saw scientific explanation as nothing more than the description of documented functional relationships. Last, Mach proposed that the only effective epistemology must be empirical, with the careful manipulation of events in controlled settings undertaken to discover more of the relationships among those events. In this last regard, Mach perceived the term *events* to be synonymous with the term *behaviors*.

With the advent of a well-defined empirical science, or research method, a flurry of early 20th century experimentation in the psychological sciences took place to further the analysis of behavior as a research method. Edward Thorndike, for example, in his single-subject work with animal intelligence, demonstrated through his famous cats in puzzle boxes experiments that animals could escape from complex mazes in shorter and shorter periods of time, documenting what we now term *trial-and-error* learning. Thorndike's puzzle box experiments were meant to demonstrate that the complex behavior of higher organisms previously thought of as largely a consciousness or mental process could be explained in terms of the inevitable result of a simple principle like the law of effect. Thorndike's work also provided a foundation for what the behavior analysis literature views as operant conditioning. *Operant conditioning* refers to the principle that future responses are changed and conditioned by the form and character of a past response.

A contemporary of Thorndike, Ivan Pavlov, through his animal research, discovered most of the facts and principles of what we know today about respondent functional relations. His empirical investigations provided the first illustrations of convincing experimental

analyses of behavior. Two main features of Pavlov's work were careful specification of independent variables (which allowed technological replication of his experiments) and attention to operational detail with measurement systems. In addition, he contributed these important scientific terms to behavior analysis: *conditioning, extinction, discrimination, generalization, unconditioned stimuli,* and *conditioned stimuli*. Pavlov documented that an unconditioned reflex could not function as the basis for more complex behavior because there would be no new functional relations and therefore no learning. Through his famous classical conditioning experiments, in which the unconditioned response of dogs salivating in the presence of food was linked with the conditioned stimulus of a bell until the dogs salivated at the sound of the bell even in the absence of food, he proved that new functional relationships between stimuli and responses could be developed and eliminated in the laboratory and therefore potentially in applied settings with humans. His contribution was, therefore, a pivotal step toward the completely mechanistic accounting for behavior that Mach proposed.

In 1913, John Watson, in his *Psychological Review* article "Psychology as the Behaviorist Views It," provided what could arguably be the most effective scientific call for a focused behavioral position in the psychological and social sciences. Watson took an extremist position that consciousness and mental processes do not have any explanatory value with regard to how humans and animals behave and interact. This position provided the conceptual foundation for a behaviorist movement in the experimental and applied sciences. Based also in large part on the logical positivism of Auguste Comte and the behaviorist views of Bertrand Russell, a formal behaviorist movement in the empirical sciences began. Out of this movement, Loeb and Crozier provided important empirical work that laid the foundation for Skinner's *The Behavior of Organisms* (1938). It is in this text that a detailed and referenced account of the brief historical summary provided above, and select referenced examples of the individuals highlighted, may be accessed for those interested in a more detailed account than that given here. Skinner's pioneering work resulted in the establishment of the Association for Behavior Analysis that thrives internationally to this day.

The contributions of Jacques Loeb are important to explain in relation to his study of the behavior of the entire organism using invertebrate animals. Loeb attempted to explain all behavior occurrences in strictly mechanistic terms. Although this attempt was not completely

successful, Loeb contributed much to a more thorough documentation of organismic function without the use of any inferences to mental functions. W. J. Crozier followed a similar quest to avoid all inferences to the mental cause of behavior, and through his work he provided the initial impetus for developing mathematical functional relations between environmental variables and behaviors of organisms. Crozier, who was a faculty mentor of Skinner's when Skinner entered graduate school at Harvard, was the primary influence on Skinner's proposed three-term contingency of Stimulus→Response→Consequence (S→R→C→), and, indirectly, the initial influence on Kantor's conceptualization of the importance of transactional relationships among behavioral events and environmental stimuli. It is primarily to Skinner, and to some extent to Kantor, that the current methodological state of behavior analysis can be attributed. Although Skinner and Kantor diverged in their emphasis of some methodological points, they both provided a fairly complete conceptual language and laid out a rigorously defined set of procedures for conducting behavioral research. In Skinner's pioneering text, *The Behavior of Organisms* (1938), he laid out almost all of the basic concepts that we rely on today, including respondent and operant procedures, rate of response as a main operant dependent variable, the cumulative record as a way of studying rate of response, many methods and results of operant conditioning and operant extinction, complete definitions of conditioned and unconditioned reinforcement, a theory of operant stimulus control as separate from conditioned elicitors of respondent functional relations, stimulus generalization theory, the idea of schedules of reinforcement, the role of motivative and emotional variables that can be defined behaviorally, and the idea that mental processes and private events operate according to the same general scientific laws that govern overt behavior. Later texts in the education and psychological sciences have provided detailed accounting of these principles in relation to applied behavior analysis as a research method and can be relied on for reference (e.g., Cooper, Heron, & Heward, 1987; Sulzer-Azaroff & Mayer, 1991).

Alone, and in collaboration with a number of colleagues—Keller, Ferster and others, to name a few—Skinner (1953, 1956, 1968, 1983, 1984, 1989) also provided a variety of methodological works related to the study of verbal behavior, schedules of reinforcement, pedagogical practices, and the science of behavior in relation to cultural change. His most notable cultural work was a utopian novel written in 1948,

Walden Two. This work was written largely in response to the more metaphysical writings of Henry David Thoreau (1962), in which explanation of the human condition was primarily couched in metaphysical and cognitive terms.

As counterpoint to Skinner's work, J. R. Kantor (1953, 1959, 1969) provided a detailed theoretical model of behavior that (a) minimized the causal emphasis of the lineal mechanics of Mach and (b) endeavored to broaden the view of behavioral interactions to a systems view founded largely on Einstein and Infeld's pioneering contributions to physics in their text *The Evolution of Physics* (1938). Kantor provided a descriptive-analytic alternative to the more narrowly focused causal determinism of Skinner. Kantor's major contribution was the recommendation to measure behavior occurrences according to the time-based sequence in which they occurred in addition to the characteristics of each behavior occurrence itself. This emphasis on the analysis of time-based sequence is most often referred to as the *probability* of a certain behavior or event following after a primary event of interest. In other words, Kantor was primarily interested in which behaviors tended to follow or precede others in time, and not simply how often or how long a particular behavior occurred. He termed his views *interbehaviorism* to communicate that a behavior must always be defined with regard to the larger time-based field of events in which it has occurred. Kantor attempted to direct methodological thinking away from lineal mechanism and the search for ultimate causes, but he was criticized for profferring a philosophy without a method for collecting and analyzing behavioral data in the ways he recommended.

With significant advances in computer technology, many contemporary behavioral methodologists have returned to a Kantorian view to provide a foundation for the development of tools capable of collecting and analyzing the types of time-based measures that Kantor originally recommended. Bakeman and Gottman (1986), for example, provide a detailed illustration of what they term *sequential behavior analysis* methods and provide a reasonably complete accounting of empirical studies using these methods in a range of social science areas, including family therapy, marital interactions, work with disabled children and youth, and educational concerns. Ray (Ray & Delprato, 1989), a contemporary of Bakeman and Gottman, provides a detailed methodological illustration of what he terms *behavioral systems analysis* in an effort to provide a thoroughgoing methodology to Kantor's

theoretical position. Sharpe and colleagues (Hawkins & Sharpe, 1992; Sharpe, 1997a; Sharpe, 2001; Sharpe, Hawkins, & Lounsbery, 1998; Sharpe & Koperwas, 2000) provide illustrations of efforts by methodologists in education to develop a thoroughgoing systems-oriented behavioral methodology and design the computer hardware and software tools to adequately support that methodology. Evidence of methodological movement toward a systemic view of the applied analysis of behavior also includes a presidential address given to the Association for Behavior Analysis by Ed Morris and printed in *The Behavior Analyst* in 1992.

In the past few decades, behavior analysis methods have frequently been used in the social and educational sciences, though they have not always been labeled as behavioral. For example, Allport (1961) pioneered what many current qualitative and ethnographic researchers currently term an *idiographic* approach (i.e., the study of the single case) as an important means of discovering information not readily available through the use of more traditional between-group designs. Doyle (1990), in recommending a direction for educational research, argued for increased effort and focus on the quantitative determinants of education practices in specific situations to more accurately discriminate the multiple event relationships among educational participants. He also recommended a behavioral approach to defining particular educational practices in order to increase the accuracy (i.e., treatment fidelity) with which these practices are implemented by others.

Currently, a variety of journals are devoted to the publication of behavior analytic studies across a variety of disciplines. The *Journal of the Experimental Analysis of Behavior*, for example, was started in 1958 to fill the need for a journal outlet for publishing laboratory-based behavior analysis activities and because of the lack of receptivity for this type of work by more visible outlets such as the *Journal of Experimental Psychology* and the *Journal of Comparative and Physiological Psychology.* The *Journal of Applied Behavior Analysis* was founded in 1968 to serve behavior analysis research in applied settings. Other journals were developed in the past decade, such as *Behavior Modification* and *Behavioral Assessment*, to provide additional reporting outlets for behavior analysis research. In addition, journals such as *Behaviour Research and Therapy* were founded in Great Britain as behavior therapy periodicals, providing articles related to respondent-based work and

some operant-derived work as well. Though not originally developed for articles devoted to behavior analysis methods, journals such as *Behavioral Science*, initiated by James Greer Miller, are now receptive to publishing behavior analysis studies and articles devoted to forwarding innovative behaviorally based methodological innovations.

In the past few decades, other discipline-specific journals either have been founded for the purpose of publishing behavior analysis research or have become receptive to behavior analysis as a legitimate methodology in the context of publishing empirical efforts. Some examples include the *Journal of Behavioral Education, Education and Treatment of Children, School Psychology Quarterly, Child Development,* and *Teacher Education and Special Education*. In addition, the Association for Behavior Analysis (ABA) was formed in 1974 to serve an interdisciplinary audience of researchers, university faculty, and a variety of professional practitioners with an interest in the experimental and applied analysis of behavior and in behavioral treatment applications. The association has steadily grown to include hosting international conferences throughout the world, financial backing to provide scholarships and grants for those performing important behavior analysis activities, and a formalized certification process for graduate-level applied behavior analysts. The ABA is *the* resource for all conference information and related professional opportunities.[1]

Thus there has been progressive interest in and a plethora of activity involving applied behavior analysis. At this point, it seems important for us to offer the following clear, concise definition of *applied behavior analysis* (Cooper et al., 1987):

> Applied behavior analysis is the science in which procedures derived from the principles of behavior are systematically applied to improve socially significant behavior to a meaningful degree and to demonstrate experimentally that the procedures employed were responsible for the improvement in behavior. (p. 14)

The remainder of this chapter presents a case for direct behavioral observation methods, summarizes a compatibility theory with respect to other research methods, and debunks some of the myths surrounding behavior analysis in the contemporary research methods literature.

SCIENCE AS A DIRECT OBSERVATION, DESCRIPTIVE ENTERPRISE

Very broadly defined, the term *science* is synonymous with the quest for knowledge (Sharpe & Hawkins, 1992a, 1992b). Again, according to Mach the term *science* implies the direct observation of things so as to produce data with a high degree of intersubjective agreement, or objectivity (Neale & Liebert, 1973). If we are to subscribe to the importance of objectivity as a necessary characteristic of science, then "doing science" includes in its definition concepts such as operational measurability, reliability and replicability of description and explanation, and systematic testability of a variety of treatments. Simply put, research in the psychological, social, and education sciences should be based on directly observable empirical evidence and reasoned argument rather than on the opinion and perspective of experimental participants (Mayer, 2000). Mayer (2001) offers two important rationales for this view of science as follows:

> 1. Science is self-correcting, so that unproductive theories eventually can be discarded on the basis of mounting evidence and reasoned argument. The alternative view that all forms of discourse—including art—are equally valid leads our field to [unbridled] relativism.
>
> 2. Rejecting science [in the traditional sense of the definition] will not garner much respect from the [larger] scientific community. Given the generally low credibility of [educational and social science] fields (Levin & O'Donnell, 2000), turning applied research into a non-scientific enterprise is unlikely to improve the situation. (p. 29)

It is in the context of such rationales that a brief epistemological overview of some important philosophical traditions is presented here to give the reader a better understanding of the importance of a direct observational approach to doing science.

How knowledge is defined and generated, and how researchers construe explanation, varies across research methods. In our judgment, researchers can benefit by a discussion of (a) the philosophical foundations from which various research methods originate, and (b) the main competing research methods approaches, including some of their main advantages and shortcomings. Table 1.1 is provided as an overview of

Table 1.1 Philosophical Traditions

Essentialism (Galileo, Newton)

1. is the strongest proponent of causal inference, of which there are three necessary conditions:
   (a) contiguity between presumed cause and effect
   (b) temporal precedence of the cause specific to the effect
   (c) constant conjunction (i.e., causal presence whenever the effect is obtained)
2. posits that causal inference assumes dependency between presumed cause and effect; a mechanistic, push-pull philosophy
3. proposes that cause does not have to precede an effect, but that the two variables must be related
4. asserts that causal inference and explanation are synonymous
5. has experimental variables explain a phenomenon in terms of their necessary and sufficient presence for the event to occur
6. assumes powerful prediction and successful forecasting
7. is highly aligned with an activity theory of causation in that it posits that causes are what we can manipulate
8. is a proponent of closed-system research, from which all known extraneous forces are excluded

Rationalism (Bunge, Descartes)

1. clearly identifies causal relationships
2. proposes that causal relationships exist outside the human mind and cannot be perceived with total accuracy by our imperfect sensory and intellectual capacities
3. asserts the existence of a special survival value to knowing about manipulable causes

Empiricism (Locke, Hume)

1. divorces the notion of causality from explanation, without abandoning the notion of cause
2. posits that all knowledge stems from experience or sensory perception, as opposed to rationalism, which proposes that knowledge may stem from sources outside of human perception

(Continued)

Table 1.1 Continued

3. asserts that knowledge is limited to what can be experienced and that what we experience is accurately and universally perceived, which places primary emphasis on our power of observation

Probabilism (Suppes, Toulmin)

1. proposes that the real cause of effects is not certain
2. posits that presumed causes have relative probability and de-emphasizes causal inference
3. asserts that presumed causes are fallible rather than inevitable
4. proposes that empirical observations can be explained in terms of an infinite number of hypotheses

Pragmatism (Dewey, James, Peirce)

1. posits that explanation that is useful in a practical sense is the only viable scientific pursuit
2. seeks to mediate between opposing philosophical perspectives
3. deplores the search for ultimate cause or absolutes
4. asserts that nature is the reality beyond which we cannot go
5. settles for a pluralistic multiple-perspective approach
6. is aligned with relativism

Positivism (Comte, Russell)

1. de-emphasizes the notion of cause in its entirety
2. posits that knowledge of cause is unnecessary because it is not implied by functional mathematical relationships
3. takes a descriptive, noncausal approach to explanation
4. asserts that correlation does not imply causation
5. emphasizes probabilism and pragmatism

Falsificationism (Popper)

1. is diametrically opposed to the descriptive "confirmationist" position of positivism

2. stresses the ambiguity of confirmation, in that corroboration gives only the comfort that the presumed cause has been tested, has survived the test, and is thus not yet disconfirmed

3. implies a skeptical probabilistic nature

4. proposes the inescapable predicament that we cannot ultimately prove a causal proposition

6. establishes cause via elimination of rival-case possibilities

Idealism Versus Realism: A Philosophical Subcategory

1. *Idealism* (Plato, Berkeley, Kant, Hocking)
   (a) couches causal inference in terms of the laws of thought, consciousness, and "hidden" meaning rather than by methods of objective science
   (b) asserts that reality consists of ideas, thoughts, minds, or selves rather than material objects and forces
   (c) posits that all we are aware of is perceptions; therefore, perceptions become reality, which is closely related to metaphysics
   (d) is aligned with phenomenalism, which involves a search for meaning behind the observable and the distinction between the appearance of reality in consciousness and reality itself
   (e) is aligned with existentialism, which involves humans' attempt to describe their existence in terms of inner conflict and its origin (e.g., Kierkegaard, Nietzsche, Sartre).
   (f) asserts that relativist and pragmatist camps are threatening to traditional science, for if the truth is relative to individual perceptions, there is no consensus

2. *Realism* (Aristotle, Macmurray, Whitehead)
   (a) posits that the state of being is existent, as opposed to that which is metaphysical or in our thoughts
   (b) asserts that the objects of our senses are real in their own right and exist independently of their being known to, perceived by, or related to mind
   (c) is most highly aligned with essentialism and causal inference due to its proposing the unchanging nature of the physical realm and the existence of universals apart from our sensory experience

SOURCE: Sharpe and Hawkins (1992a), based on Bakker and Clark (1988), Cook and Campbell (1979), and Titus, Smith, and Nolan (1986).

philosophical and epistemological traditions to guide discussion. Although there may be disagreement among scholars regarding some of the finer points in the table, Table 1.1 provides a good general introductory summary to the philosophical traditions.

It is important for researchers considering a particular research method to understand that method in the contexts of both other methods and the philosophies that guide the method (compared to the philosophies that guide other methods). Integrally related to philosophical understanding is epistemological familiarity. In other words, when a researcher chooses a method of data collection and analysis, what is really chosen is the explanatory adequacy of a particular set of procedures. "Doing science," then, refers to the investigation of events by purposeful and strategic observation. A philosophy of science concerns the study of science itself, or the question of the relative utility and legitimacy of a particular research method. We therefore have a continuum of philosophy, epistemology, scientific activity, and ultimately the application of scientific results in professional and cultural settings. For the purpose of this discussion, we will look at the elements on the continuum in reverse order. At one end of the continuum is a technology of application, which does not test scientific propositions but assumes them in order to design practical applications to benefit certain individuals or groups. Adjacent to technology on the continuum is science, which builds knowledge and information structures based on the systematic observation and analysis of data. Philosophy and epistemology are last on the continuum and concern themselves with the legitimacy of the scientific method and the logic of that method's scientific practice and scientific explanations.

We argue within a behavior analysis framework that direct observation, description, and the resultant explanation are all necessary and central components of doing science. A traditional definition of *explanation*, based in essentialism and rationalism philosophies (refer to Table 1.1), has been that it is what most clearly identifies a cause for something. Applied social and psychological scientists throughout the 20th century routinely and predominantly employed research methods aligned with these two philosophical perspectives. This causal perspective is also integral to Skinner's behaviorally based three-term contingency model.

However, current methodological debate in the social, educational, and psychological sciences centers around the supposed inappropriateness of adopting the types of philosophies and related research

methods just summarized. One issue being debated is the notion that the strategies and tactics used to generate knowledge from the nonhuman or laboratory culture of science ought to be somehow different from the strategies and tactics used to generate knowledge about human interactions in applied settings. While we bring into question adopting an essentialist or rationalist philosophy as the most appropriate foundation for engaging in scientific activity, we don't do so by arguing that studying applied settings and human interaction is somehow fundamentally different from studying nonhuman phenomena in controlled laboratory settings. At issue for us is the importance of understanding the foundational philosophies that guide particular scientific methods, and in particular those that guide Skinnerian and Kantorian behavior analytic methods in relationship to those that guide other potentially compatible methods of data collection and analysis.

For example, leading empiricists such as Hume and Locke argued against a purely causal view of science as a viable construct (Hendel, 1963). Those who have accepted this empiricist position tend to deemphasize the search for independent cause as the appropriate model for explanation. A more extreme position has been taken by some philosophers who are generally associated with the positivist school and who doubt our ability to ever know the causes of events with any certainty and therefore advocate that scientists minimize or even abandon the notion of causality altogether (Russell, 1929). Instead, simple description of events of interest as they naturally occur is equated with explanation, with any reference to one event causing the occurrence of another removed from the scientific enterprise.

We argue, however, that the total rejection of causality is unwarranted. Like Suppes (1970) and Toulmin (1961), we promote the idea that causal connections between directly observed events are more accurately characterized as "probabilistic." The methods provided in this book are aligned with Cook and Campbell's (1979) conception of cause that avoids an essentialist or purest explanation and instead recommends the scientific probing for high probability causal connections, using description (e.g., visual inspection of data graphs) as the primary data analysis vehicle.

Essentially, we promote an empirical view of research that deemphasizes, but does not entirely reject, the notion of cause. Though most philosophies respect the idea that scientific activity is designed to explain human experience, essentialism and rationalism identify such

explanation with the location of causes, while empiricism tends to divorce the notion of causality from explanation without completely abandoning a causal conception. We argue that scientists should perhaps take more seriously the proposition that all we know scientifically we learn from direct observational experience, and couple it with the judgment that experience tells us "that" something is the case rather than "how" or "why" it is the case (Hendel, 1963). This philosophical position leads to the intriguing possibility that explanation is very similar (if not identical) to "description." If we are to agree that explanation is constantly evolving toward more accurate scientific descriptions of the world around us, and that the most plausible explanation of an event at a past point in time may prove to be erroneous at present and even change again in the future, then perhaps the best we may hope to achieve scientifically is accurate, all-inclusive description of our world. Science, then, can be argued as a largely descriptive enterprise, with researchers primarily engaged in reporting and organizing objectively observed events.

A final proposition that we take seriously, and one that is a fundamental component of behavior analysis methods, is that an adequate scientific description/explanation is one that is deemed useful and beneficial, or pragmatic, to a particular individual or group. Especially important to education and clinical psychology practice is the idea that the best scientific theories are those that offer practical or therapeutic advantages. Many clinicians in the psychology profession have repeatedly voiced that much of the research available is of little use in guiding effective clinical practice (Bergin & Strupp, 1972). Part of the problem with the research lies in the preponderant use of large-group comparison methods, with assumptions of ultimate cause, rather than methods more amenable to focusing on challenges with individuals and documenting specific strategies that may be of therapeutic benefit. Typically, the practicing clinician or educator is confronted with challenging situations at the individual level (Kazdin, 1982), and "it is at that level that empirical evaluations of treatment need to be made" (p. 14). With regard to education, the most scientific explanation in teacher effectiveness research is the one that helps us best integrate a teacher's intent; the content being taught; and the context, process, and outcomes of particular teaching practices; it also assists in looking for ways to improve all of these components (Metzler, 1989). The focus of this book, then, is to provide contemporary research methods in a behavior analysis framework that are probabilistic, empirical, and

pragmatic as well as deemed helpful to understanding more of what may be therapeutic for particular individuals in applied settings.

TOWARD A MULTIPLE RESEARCH METHOD COMPATIBILITY THEORY

Although this book is designed to explain how to perform behavior analysis research, and we advocate its use for a variety of research and evaluation activities, we also recommend that behavior analysis methods be used in concert with other research methodologies. In other words, although we explain and recommend one method, we feel that it should not be used to the exclusion of, or independently of, others and that other methods may benefit from the addition of a behavior analysis component. Clearly, this is an important topic for those who are being trained in the activity of engaging in applied science. If this position is adopted, then researchers should be supported by being adequately armed with all of the data collection and analysis tools necessary to match specific research questions with the most appropriate combination of scientific practices.

Support for most applied education and social science research activities has not always been forthcoming from the basic or experimental laboratory science community. Indeed, the basic or laboratory science community often views many of the published forms of education and social science in applied settings with great skepticism. In some ways, this skepticism is the result of the difference between what the basic science community thinks of as "art" and what it views to be genuine "science." Long marginalized by mainstream laboratory scientists, researchers in education and social science, and to an extent the psychological science, have wrestled with the challenge of how to accurately and inclusively describe and analyze applied settings in a quantitative way. This has been particularly the case with those settings in which a host of complex behavior-behavior and behavior-environment relationships are ongoing with high frequency and in simultaneous or overlapping fashion.

Simplistically put, there have historically existed the two broad categories of qualitative and quantitative research methods in the social, educational, and psychological sciences designed to meet these challenges (Shulman, 1987). In the most radical case, some (e.g., Barone, 2001) have seriously proposed that applied education research in

particular should become nonscientific, so that, for example, artistic productions would be considered to be educational research studies. The argument goes something like this: By turning away from traditional quantitative science, education researchers would become unfettered by the annoying requirement to base arguments on directly observable evidence. Relativism would be the foundational philosophy (refer to Table 1.1), in which all individual opinions become equally valid explanations for phenomena. We argue, as does Kerlinger (1986), that opinion and authority are the most problematic and flawed ways of gaining information of value. However, we also argue that there are a range of research methodologies that are compatible with the applied analysis of behavior. We discuss these methodologies here from a compatibility perspective to foreground the remaining chapters of this text.

The two general categories of research methods have been more often opposed to each other than seen in terms of their potential compatibility. One category is qualitative research, of which ethnographic narratives and self-report questionnaire-type efforts have generated the most interest and are the most well-known examples. This family of research draws heavily from anthropology, sociology, and linguistics. It relies on an interpretive focus and utilizes a largely subjective perspective (Smith & Lytle, 1990). The other general category is quantitative research, which includes behavior analysis as well as a variety of descriptive, correlational, and experimental methods.[2] Discussion both in the contemporary applied literatures and at professional conferences has been devoted to the rhetoric of qualitative versus quantitative research; more typically than not promoting the incompatibility of each category due to intractable differences in their underlying philosophies (Firestone, 1987).

Two dominant systems of thought in the education, social, and psychological sciences—social constructivism and radical behaviorism (subsets of the larger qualitative and quantitative research methods categories, respectively)—have tried to meet the challenge of combining qualitative and quantitative methods and have consequently provoked a degree of resistance within the larger scientific community and have been subjects of attack (see Binder, 1994; Brown, 1980; Glaser & Strauss, 1967; and Morris, 1984, 1992 for a more complete discussion of these issues). In addition, Newman's (1992, p. 13) point remains the case—"proponents of . . . these [two] viewpoints have rarely accepted the other as valid, and have been at philosophical war with one another" since their respective inception. We agree with Newman

when he goes on to say that these two methodologies contain more common characteristics than differences and that perhaps each could help with the often cited criticisms of the other. Also, behavior analysis methods, as a subset of quantitative research methods, may prove to be a bridge across the two larger categories of qualitative and quantitative research.

Qualitative Research

Most qualitative research stems from the field of cultural anthropology. In this method, the research setting is a catalyst for stimulating interest in the researcher in inductive ways. The researcher is an active participant and an integral part of the study. The focus has typically been upon the relationships between the environmental events that occur within natural settings and relevant participant responses. A rigorous set of scientific procedures has been developed to collect and analyze narrative descriptions of a variety of natural settings. A similar set of procedures has also been developed to collect and analyze various types of interview responses designed to gain insight into the participant perspectives and cognitions believed to be operating within those settings (see LeCompte & Preissle, 1993, and Miles & Huberman, 1984, for a detailed discussion of such methodological procedures). This research category provides a potentially inclusive description of behavior-environment events, and potentially provides data related to thinking and mind, getting at what many behavior analysts term *private events* or *covert behavior*.

A lot of confusion, however, surrounds those who discuss qualitative research, in that it has often been regarded as if it were one approach (Lutz & Ramsey, 1974; Magoon, 1977; Rist, 1977; Smith, 1983; Wilson, 1977). Jacob (1988) and Smith (1987) provide detailed discussion of qualitative research methods across a broad classification of types and strategies and into the general areas of: (a) human ethology, (b) ecological psychology, (c) holistic ethnography, (d) cognitive anthropology, (e) ethnography of communication, and (f) symbolic interactionism. Important to this discussion is that the central theme of all of these qualitative strategies is participant observation of as many behavioral and setting variables of interest as possible. The desired database for all of these strategies is a synthesis of extensive narrative fieldnotes. For qualitative methods, however, participant observation may be, at the same time, a primary advantage and a major shortcoming.

Due to the strategy of narrative synthesis with a single case or small group, qualitative research involves a highly reductive data analysis based in large part on researcher intuition, with conclusions couched within the inherently value-laden perspectives of the researcher and the study participants. In other words, the final data analysis is in large part dependent on the subjective way in which certain data are gathered and used while other data are overlooked or discarded during the data synthesis process. Although qualitative methods may provide important insight into the form and character of an experimental setting, direction regarding what should be studied, and participant (or social) validation of the researcher's perspective on a particular treatment, the nature of the data collection and analysis process makes it inherently difficult to distinguish cause from effect, with overall experimental coherence becoming a challenging determination.

The main weakness of this method, therefore, lies in the need for grounding the descriptive narrative data collection efforts in research techniques that can provide a more objective means of data analysis. In essence, the qualitative descriptive details from which the intuitive conclusions of the researcher stem are in need of criteria for consistent data interpretation across studies.[3]

Quantitative Research

Quantitative research has a longer formal history than its qualitative counterpart and consists of many forms, including descriptive, correlational, and experimental methods and the behavior analytic procedures that are the focus of this text. The chief advantages common to this variety of data collection and analysis strategies are a more objective data collection and analysis process and a clearer demonstration of causal connections or potential functional relations among variables. This category of research methods, however, is not free of methodological challenges either, although many myths and misconceptions have been cast upon these methods during contemporary methodological debates in the literature (see the Some Myths About Behavior Analysis Debunked below for more on this). Historically, quantitative research in the applied social, educational, and psychological sciences has passed through overlapping phases ranging from efforts to simply describe existing states to correlational activity among behaviors and events within a particular setting to large-group comparisons across distinct participant groups that each receive different

types of treatments. An accompanying array of parametric and nonparametric statistical analyses of numerical data have developed in complexity and sophistication over the years to accommodate the needs of quantitative data analysis. Although behavior analysis has not received the same legitimacy in mainstream quantitative methodological circles as other forms of quantitative research, it is included in this research category due to its thoroughgoing quantitative character, and even though it also employs many characteristics of qualitative methods (including intensive study of individuals or small groups and a focus on the direct observation of the behavior and environment events that occur in the settings in which certain individuals operate). In this regard, behavior analysis may be viewed as a compatible or collaborative bridge across quantitative and qualitative research methods.

Behavior analysis is a more focused avenue of study than its qualitative counterpart. In similar fashion to most quantitative methods, however, by stipulating quantitative behavior and event category systems prior to data collection, researchers may to some extent be imposing their own ideas on a setting, because data collection is preceded by knowledge of what a researcher wishes to observe. One of the most pressing traditional constraints to quantitative and behavior analysis strategies lies in the practice of isolating behaviors and events from the complex stream or chain of events in which they naturally occur in applied settings. This is a particular challenge to behavior analysis methods in light of the claim that behavior analysis is aligned with the philosophical school of *contextualism* (i.e., the view that all behavior and event activity must be studied with sensitivity to the particular context and setting within which it occurs; Morris, 1992). What will become apparent in later chapters is the challenge that more Skinnerian and lineal mechanic behavior analysis methods have been faced with when certain behaviors and events under study are removed from the larger context of an experimental setting due to the method of data collection and consequent analysis. The response to this is the sequential and interbehavioral behavior analysis methods that will be discussed in detail in later sections of this text. These strategies are specifically designed to take into account more inclusive description of multiple characteristics of multiple behaviors and events, and detail in quantitative ways the time-based connections among behaviors and events as they actually occur in an experimental setting.

Another principle that challenges the legitimacy of the logical positivist philosophy on which most quantitative methods are based arises

out of the dilemma concerning the quantitative methodological process of confirming scientific theories (Garrison, 1986). Confirmation of anything from grand theoretical statements (T) to the most basic hypotheses (H) consists of drawing a logical or causal implication from the hypothesis or theory to an experimental conclusion (E). In the quantitative research tradition, this is stated using the following structure: If T or H is true, then E will be observed; and, consequently, if T or H is not true, than E will not be observed. If the experimental conclusion is observed under the appropriate research conditions often enough to be statistically significant (i.e., occurring more often than would be assumed by random chance) through use of an acceptable statistical analysis technique, then the theory or hypothesis that is the focus of a study is said to be verified. The problem is that this structure is logically flawed, representing what has been called *the fallacy of affirming the consequent* (Garrison, 1986; Johnston & Pennypacker, 1980). For example, a particular theory or hypothesis may be true and the experimental conclusion nonetheless false due to other explanations and impacting variables. In contrast, a particular theory or hypothesis may be false but supported as true due to experimental conclusions that are connected not to the theory or hypothesis but to some other variable or group of variables that was responsible for the experimental result that was recorded. This provides the inherent danger to quantitative research of statistically confirming a variety of theories and hypotheses through inappropriately connected experimental conclusions and the resultant explanations.

Methodological Compatibility

Although both general research categories that we have presented provide clear advantages for particular research questions and related research study designs, some challenges to each may be outlined as follows:

1. Qualitative research is
   (a) extremely broad in focus, producing a rich, descriptive database. The database, however, may be value laden as a function of how a researcher collects the data, which in turn impedes the probability of a completely objective data synthesis and analysis process when reporting research results.

(b) founded philosophically in relativism and phenomenology, encouraging researchers to ascribe larger meanings to behavior and event occurrences. The objectivity of the data is not checked, resulting in an oftentimes unbridled subjective and intuitive data analysis process.

2. Quantitative research is
   (a) explicitly objective in data collection and analysis; however, many time-based interactions among variables in an experimental setting are oftentimes ignored or excluded through methods intended to simply control for the potential effects of variables not under direct experimental analysis.
   (b) a largely deductive process with regard to data analysis, with the presupposed methodological structure and procedures of a particular experiment inappropriately providing untoward emphasis on the ultimate conclusions derived from a data analysis of experimental conclusions.

Again, debate currently centers around the perceived necessity of qualitative and quantitative research methodologies remaining separate due to their supposed incompatibility (Howe, 1988). The main argument in this debate centers around how closely particular research methods must be tied to their respective philosophical foundations. Firestone (1987), for example, provides an appealing argument that philosophical ties are rhetorical at best. We know that quantitative methods express the assumptions of positivist philosophy, which holds that behavior can be explained through objective data. In quantitative methods, design and instrumentation persuade by showing how bias and error are eliminated. On the other hand, qualitative methods express the assumptions of phenomenological philosophy, which states that there are multiple realities that are all socially and perceptually defined in a variety of ways by particular individuals and groups. Rich description persuades through a researcher, who, having been immersed in a setting of interest, provides enough verbal detail to enable an audience to make intuitive sense of the situation. Firestone (1987) proposed that although the two methods are rhetorically different, their results can be complementary and therefore compatible.

Finn (1988) and Shavelson and Berliner (1988) emphasized the need for using alternative combinations of research strategies in the social and educational sciences that draw from the advantages of each

particular research method. For example, quantitative research methods may count up the number and duration of deviant behaviors in a classroom situation and find that certain teacher practices diminish those behaviors, whereas qualitative research methods may discover some of the reasons behind the incidence of that deviant behavior and the relative receptivity on the part of the teacher to implement practices designed to reduce the unwanted behavior. Finn, as well as Shavelson and Berliner, stress that the major ailment of applied research in general is the public's great skepticism toward it. Although education and social science research, and even much psychological research, yearn for the attention and respect given the laboratory-based physical sciences, the physical sciences have gained respect due to the enormous practical difference they have made through their technological application in the lives of people. We argue that in some respects the lack of public respect for education and social science research is due to a lack of research methods collaboration, and that this lack is related to the lack of impact of separate efforts on the public (and professional) good. In essence, the lack of research methods collaboration and a proactive and supportive research community may serve to limit the fruitfulness of education and social science research in practical terms.

During the past three decades, the scientific community's opinion of qualitative methods in educational research has evolved from ridicule to appreciating its utility in provisional exploration to wholehearted acceptance of it as a valuable alternative in its own right to seeing it as a method that should receive priority in terms of scientific legitimacy. We argue that what is necessary is a return to perceiving all research methods as equally important in the context of the types of questions that they are best suited to answer, and that qualitative methods should be viewed as potentially compatible with quantitative strategies.

A major concern of those who argue for incompatibility is that capitulation to what works experimentally ignores the supposed intractable incompatibility of the competing realist and idealist philosophical outlooks (refer to Table 1.1) that support quantitative and qualitative methods, respectively. However, Howe (1988) and Newman (1992) propose an appealing pragmatic alternative. They argue that no incompatibility between quantitative and qualitative methods exists at either the level of practice or the level of epistemology. Taking a pragmatic stance, they further argue that there are no reasons at the level of scientific activity for applied research to resist

forging ahead with what works due to the many commonalities of qualitative and quantitative research at the level of data collection and analysis—in particular with regard to behavior analysis and qualitative methods.

We feel that research methods compatibility rests upon two points. The first point involves the domain of research practice. In practice, differences between quantitative and qualitative data, design, analysis, and interpretation may be accounted for largely in terms of differences in research interests and judgments about how best to pursue them. For example, if a teacher wants to know how many correct responses students are making with respect to a particular method of teaching reading pronunciation, then a quantitative data collection method is most appropriate. If the same teacher, on the other hand, wants to know how receptive his or her students are to the type of instruction being used, then a qualitative data collection method may be more appropriate. If differences can be accounted for by research interests and how best to pursue them, it prompts suspicion about the need to promote different conceptions of reality and different philosophical underpinnings as the rationalization for using different research methods.

Our second and related point is a bit more elaborate and depends on acceptance of the first point. Researchers who argue for intractable incompatibility admit that problems arise not so much at the level of scientific practice as at the philosophical level via the following rationale: Realist and idealist philosophies underlie respective quantitative and qualitative methods. These two philosophies are purported to be incompatible due to their contradicting premises. Therefore, the two methods are incompatible. Our response to this is that a principle implicit to incompatibility argument—that abstract philosophy should determine research methods in a one-way fashion—is erroneous. Research methods, in our view (and also implicit in a behavior analytic view), must demonstrate their worth in terms of how they inform, and are informed by, their respective audiences and clientele to potentially therapeutic ends. If such a two-way relationship is viable, then research methods, and potential combinations thereof, must be evaluated in terms of how well they correspond with the demands of research practice, and, if warranted, the idea of the intractable incompatibility of methods must vanish.

Although each research method has a different philosophical foundation and different and unique data collection and analysis

procedures, each also has particular benefits that may complement the other. Due to the ability of qualitative and quantitative methods to each potentially negotiate the obstacles of the other, alternative strategies that incorporate some of the benefits of both research categories may be valuable (refer to Jacob, 1982, for a more complete discussion of this). Bliss, Monk, and Ogborn (1983) provide an early example of a categorical quantification system used in qualitative research and designed for methodological compatibility. Experimental efforts required the coding of direct behavioral observations into categories while still preserving the inductive nature of the data. Data analysis took the form of a tree diagram, a pictorial representation that showed related, independent, and conditional categories. Such a tree structure provided an important aid in organizing large sets of descriptive categories into a structured quantitative system as well as in enhancing the clarity of qualitative research reporting and interpretation. Brown's (1980) Q-Methodology provides another specific illustration of the prospects of coupling qualitative methods' ability to uncover subtle influences in a setting context with a quantitative structure to more objectively determine behavior-environment relationships. There also exist a host of contemporary published applied behavior analysis studies that rely on qualitative techniques for socially validating the quantitative support for a particular educational treatment (e.g., Sharpe, Lounsbery, Golden, & Deibler, 1999).

Some Professional Recommendations

We believe, as do Heward and Cooper (1992) and Sage (1989), that the lack of equal and thoroughgoing exposure of researchers to a variety of research methods is due in large part to the narrow focus of traditional research methods courses in graduate programs in the educational, psychological, and social sciences. Most typically, graduate students are taught a very traditional quantitative version of knowledge production, or, alternatively, a primarily qualitative approach to science, as the only legitimate way to engage in research. If alternative forms of applied research are mentioned at all, they are more often than not presented in a less than desirable light. Instead, a convincing case needs to be made in these courses that there is a broad array of viable forms of scientific inquiry. We caution, as have others (e.g., Schutz, 1989), that strong advocacy of one particular research method, to the overshadowing or exclusion of others, only encourages new

researchers to adopt the "have method, need problem system" (p. 31). Of utmost importance is the need for the careful and reflective development of the research question prior to framing it within an appropriate data collection and analysis method. The most prominent interbehavioral theorist in psychology, Kantor (1979), maintained that any science that insists on the priority of a particular research methodology is not yet true science. In our view, that the research question ought to dictate the method of investigation is perhaps the most self-evident, yet most often ignored, principle of science. It is in this context that the methodology materials are compiled in this text in what we hope forms a resource for those interested in the principles and practice of applied behavior analysis and quantitative approaches to direct observation of behavior-event relationships. Prior to jumping into a detailed recipe for these methods, however, we believe that it is important for us to discuss some of the prevalent myths that surround behavior analysis methods.

SOME MYTHS ABOUT BEHAVIOR ANALYSIS DEBUNKED

The following statement by Silverman (1996) provides a pointed illustration of how most professionals who operate outside the university setting—those the university-based research information is designed to reach—view the research process: "The best way to have people avoid you at a cocktail party is to tell them you teach statistics and research methods" (p. 36). If this quote is any indication, a significant challenge of bridging the gap between research information and its professional application remains great and is potentially becoming worse. In addition, this challenge appears specific to the social and educational sciences. Landrum (1997) illustrates this by showing that physical science data is generally trusted; for example, when boarding an airplane there is typically inherent trust that reams of experimental data have been consulted in the construction and flying practices of each airplane to ensure optimal safety and effectiveness. When a physician is consulted by a patient, she or he will often support a proposed treatment plan by referring to the latest medical research. However, when teachers and clinicians in the social sciences are consulted, they rarely provide descriptions of research supporting a particular practice used by a teacher or social worker.

There are many potential reasons for this continuing challenge to the scientific professionalism of those in education and the social sciences. First, data tend to be ignored, with many professional practices contradicting data-supported effective procedures. Kauffman (1996) provides an example in mathematics education in which the currently popular constructivist approach to teaching math skills has been widely adopted despite clear evidence that "it is not the most effective and efficient instructional approach for most students" (p. 56). This is evidence of a larger cultural challenge, again according to Landrum (1997), which is our society's general trend of belief that applied scientific data do not matter in terms of providing any guidance to educational and clinical practices.

There are some items that we feel it is important to discuss here in relation to the pervasive nonuse of scientific data in making professional decisions. These involve the potential reasons for the nonuse of data, in general and in relation to some of the appealing features of behavior analysis research methods, and some pervasive myths concerning applied behavior analysis that are detrimental to acceptance of the method as a legitimate partner in the community of applied scientific methods.

One potential reason for the nonuse of data is that American society tends to place great distrust in information presented in quantitative formats. This is probably due in large part to the knowledge that our news and marketing media consistently manipulate information in ways designed to meet oftentimes self-serving ends antithetical to an increased quality of life. In addition, a traditional mechanistic causal model for conducting research may not be the most appropriate or capable method to use in determining effective educational or clinical practice due to the host of variables that are operating in applied instructional settings. In education, isolating one independent variable, such as teacher feedback, and selecting one dependent measure for its relative effect, such as appropriate subject-matter engagement, without taking into account the myriad other variables that may also have an effect on subject-matter engagement, will often provide for misleading, albeit data-supported, conclusions (as mentioned earlier in our discussion of theories and hypotheses not necessarily being appropriately connected to experimental conclusions).

Another challenge to the use of data-supported professional practices is that there may simply be too much ongoing data generation for professionals to feasibly keep abreast of it. This is due to recent

advances in technology that have facilitated data access, coupled with the basic governing principle for the relative success of research faculty professionals to produce as much data-based publication activity as possible. This ever-increasing activity means that there is an overwhelming amount of information to sift through. This abundance of data, and ready access to this data, provides the professional with the added challenge of reviewing and determining sound data from that produced simply for the sake of research faculty advancement and without scientific merit.

A third challenge is the dangerous assumption promoted by currently popular qualitative methodologies that there are no universal truths to be discovered, that individuals construct their own realities and beliefs, and that each of these variable individual constructions is equally valid and true. If we take this approach to applied research (and we by no means wish this challenge to be interpreted as a criticism of qualitative data collection and analysis methods, for we believe these methods to be legitimate and important contributors to the applied scientific literature), then as research faculty professionals we may be inadvertently dismissing the importance of adherence to objectively supported professional practice. In other words, if one does not strive for the data-supported discerning of static and generalized truths, then no one professional method of doing anything can be argued as a relatively more or less effective method, and therefore data collected in a particular situation with particular individuals and on a particular professional practice will have little relevance for any other situation.

We believe that using the principles of applied behavior analysis based on sound behavior analysis research is one means of overcoming the nonuse of scientific data. For this to occur, however, it is necessary for researchers to combat some popularized negative opinions about behavior analysis as a viable research method. Negative opinion can be countered with a variety of arguments.

A first and important assumption of behavior analysis research is that it must have a therapeutic or utilitarian criterion. All information to be gathered must be designed to help improve the quality of life of the participants in positive ways, and ideally all research activity should be conducted collaboratively with groups of interagency professionals, including faculty-level researchers and the professional practitioners who operate in the situations to be studied (Sharpe, 2001; Sharpe, Lounsbery, & Templin, 1997). Pragmatic questions, such as

how we want professional situations to function and how we can get them to be that way, are of paramount importance in such research.

Although behavior analysis is of potential appeal because of its pragmatic stance, and many researchers have called for greater use of data-supported principles in professional practice over the past two decades (e.g., Hrycaiko & Martin, 1996), many myths and misconceptions have halted the method's optimal impact in the education, social, and psychological sciences. These misconceptions can be categorized into five general areas (see Aeschleman, 1991, and Kazdin, 1982, for a detailed discussion of this): behavior analysis lacks internal validity, behavior analysis lacks external validity, behavior analysis uses a visual inspection of data that is not scientific, behavior analysis does not take cognitions into account, and behavior analysis is a technocratic and robotic enterprise. Each of these myths is discussed below.

Myth 1: Behavior Analysis Lacks Internal Validity

Behavioral methodologies have evolved into a complex offering of designs arranged to specifically control for most traditional internal validity concerns. These range from multiple baseline designs, in which treatment is provided to different matched participants at different points in time, to multiple treatment and component treatment reversal designs, in which treatments are provided in different sequences to different participants, to multiple treatment introductions and extended maintenance and nontreatment phases, to a variety of multiple-measure, multiple-variable alternating treatment designs (all of which are detailed in later chapters of this text in relation to validity concerns).

Myth 2: Behavior Analysis Lacks External Validity

Behavior analysis methods operate on the premise that scientific study should hold as most important the discovery of how professional participants and the environment in which they operate may be arranged to be most therapeutic for the various participants in that specific environment. Thus, behavior analysis methods emphasize discovering what might help a variety of professionals and their clientele in particular situations and with particular challenges. This is effected through careful replication of an original study that includes a potentially effective treatment, and by careful manipulation of the

characteristics of participants and environment from study to study, making rigorous generalizability of findings possible over the course of a series of focused studies.

Myth 3: Behavior Analysis Uses a Visual Inspection of Data That Is Not Scientific

This misunderstanding is potentially the most damaging to behavior analysis methods because traditional statistical analyses have a subjective warrant of translating into *good research* by the very function of their use in the eyes of many researchers and research consumers. In this regard, then, visual inspection of data represented in graphic form is considered an illegitimate form of analysis (Huitema, 1986). A contrasting argument is that all that a large group statistical comparison may hope to show through disproving a null hypothesis based on a traditional statistical manipulation is that "something" differed across experimental groups during a particular study. What a statistical analysis does not show, and what is a great leap of faith on the part of researchers using such analysis, is that this difference is actually and explicitly due to whatever treatment or independent variable was introduced to a particular experimental group. In addition, and little understood by those unfamiliar with behavior analysis methods, a sophisticated set of visual inspection practices and procedures are required to argue favorably for a treatment effect, including mean and level magnitude of change analyses, trend and latency rate of change analyses, and a variety of appropriate nonparametric statistically based data comparisons. Also, visual inspection of readily discernable data changes contained on a graph is viewed by the behavior analyst as beneficial, especially in relation to the challenge of more complex and less discernable statistical information. Changes due to a particular treatment, for example, may be more receptively argued for use by practicing professionals as a function of the relative ease with which the data may be accessed and interpreted from a graphic representation. Barlow (1980) argues that the gap between important data-supported information and its use by professional practitioners will not be narrowed as long as researchers continue to slavishly insist on traditional large-group factorial designs, multivariate statistics, and predetermined levels of significance, for these types of information will never be viewed with receptivity by professional practitioners in real-world settings.

Myth 4: Behavior Analysis Does Not Take Cognitions Into Account

Another argument against the use of applied behavior analysis methods is that the data collection and analysis procedures implemented do not take into account what experimental participants think, feel, or perceive in relation to the behaviors and events documented by the method. In response, Skinner (1945) provided the concept of what behavior analysts term *private events* (i.e., events that take place inside people and are not detectable by direct observation). From a theory of behavior perspective, both private and publicly observable events operate according to the same Stimulus→Response→Consequence laws, with the difference being one of the relative feasibility of data collection and analysis of those respective events. Although mental explanations are de-emphasized in behavior analysis methods due to the challenge of collecting data on private events, the use of questionnaire and self- report data is recommended in concert with public event data to provide supporting information on private events (these are typically called *social validation* procedures).

Myth 5: Behavior Analysis Is a Technocratic and Robotic Enterprise

The applied analysis of behavior has been traditionally thought of as a very mechanistic and rule-governed approach to collecting information, in which two or three behaviors are isolated for analysis from a larger experimental situation. In this traditional model, a relationship between two behaviors is hypothesized in a general sense, data are collected in a particular situation to support that relationship, and then general rules are promoted to guide the instruction and application of the behavioral relationship to be supported. In education, connections are made among these fragmented behaviors, in which isolated teacher events, which presumably affect student practices or student achievement in some mechanistic or additive way, are aggregated over time (Doyle, 1990). As will be seen in our detailed description of behavior analysis methods as evolved and unpacked, many computer-supported comprehensive description and sequential analysis methods have been designed and developed to discover the multiple relationships among multiple teacher and student behaviors and events in particular situations (see Morris, 1992, for a detailed discussion of emphasis on behavioral "discovery" and "understanding"

rather than simple behavioral "demonstration"). More sophisticated behavior analysis methods have been developed as a function of researchers requiring more sophisticated tools capable of discovering the multievent conditions that characterize effective and not-so-effective professional practice in particular situations and of providing information to alter those conditions and encourage those professionals in more effective practice in the context of challenging client behavior.

OVERVIEW OF BEHAVIOR ANALYSIS PRINCIPLES AND PRACTICE

The remaining two chapters contained in the first part of this book provide an important introduction to the main principles of applied behavior analysis, with the remainder of the book detailing an introductory to intermediate-level cookbook approach to the activity of engaging in applied behavior analysis research in a compilation of most of its many methodological and procedural forms. The order of information is structured to facilitate the reader's familiarity with the steps necessary to conduct behavior analysis research that would logically be taken when engaged in the actual research process. As the preface detailed, the remaining chapters in Part I focus on some of the basic principles and terminology necessary to the navigation of applied behavior analysis methods, and detail some of the important specific principles of contemporary systems-oriented and sequential methods. Part II of this book provides a detailed procedural primer for constructing a coding scheme for particular research or assessment purposes, including the many assumptions and limitations that should be taken into consideration when conducting behavior analysis research. Reliability and treatment fidelity issues and procedures are discussed in detail, and close attention is paid to the steps of criterion standard development, staff training, interobserver reliability, and treatment implementation accuracy. A series of category system illustrations taken from a variety of education, social science, and psychology disciplines provides the reader with hands-on familiarity as to how behavior analysis efforts in these respective areas has been implemented with success. Part III of this text includes a variety of reported and generally accepted techniques in the areas of collecting, analyzing, and visually representing data. Application procedure details and their potential advantages are provided and related to some contemporary techniques

for recording in real time, overcoming validity challenges through more sophisticated research design types, and issues regarding graph preparation and the use of statistical analysis support for descriptive and inferential purposes. The last part of this text provides recommendations on how applied behavior analysis methods may be used to enhance a variety of research and development, professional or clinical assessment, and instructional applications across a variety of education, social science, and psychological activities. We have included detailed illustrations of field-based professional evaluation activities, research and development opportunities designed to uncover information not previously available to other research methodologies, and laboratory simulation activities heretofore unavailable through other methods and without the aid of behavior analysis-supported computer technology.

As computer hardware development moves forward with increasing speed and sophistication, our potential for as yet unrealized quantitative description and analysis of the world around us is greatly enhanced. Those of us engaged in the development of increasingly capable observational instruments are continually amazed at not only the wealth of additional information that is becoming available but also the opportunity afforded to look at things in new and different ways. In these contexts, we want to provide researchers, educators, and clinicians with a quantitative means to more capable and inclusive description of highly complex and interactive settings. In compatible software publications (our BEST software, described in Appendix B and advertised in the flyer included with this book; Sharpe & Koperwas, 2000), we provide tools capable not only of such a task but also of being programmed for the unique observational interests of a host of professionals working within a wide variety of situations and settings. To provide a theoretical foundation for these materials, we draw on the conceptual work of pioneers in observational and field theory for support for some of the alternative lenses we discuss that can be used to view and analyze the rich and varied information that state-of-the-art computer-based data collection tools are capable of providing. Finally, we include conceptual and procedural information that may help those interested in the direct observation of complex phenomena as they undertake the challenges of such an endeavor.

NOTES

1. The Association for Behavior Analysis (ABA) is currently located near Western Michigan University and can be contacted as follows:

ABA Offices
1219 South Park Street
Kalamazoo, MI 49001
Telephone: (269) 492-9310
Fax: (269) 492-9316
E-mail: Mail@abainternational.org
The ABA website is http://www.abainternational.org

2. It can be argued that even quantitative research that has been characterized as psychometric or cognitive can be behavior analytic in framework. The behaviors or the products of behavior of the experimental research participants are typically reported or inferred using these methods, though the results of study are typically interpreted in nonbehavioral terms using nonbehavioral approaches to data analysis.

3. It is important to note that qualitative research methodologists do not regard the narrative descriptive process as a weakness. They believe that the bias, subjectivity, and relativity contained within qualitative databases are meant to be encouraged and exploited rather than restrained through a mechanism for objectivity. Intersubjective agreement or data objectivity is viewed as irrelevant. Our intention is not to denigrate these views, but rather, to point out the potential drawbacks and challenges to this approach and to contend that a rejection of intersubjective agreement and objectivity as a necessary component to scientific practice is to potentially alter the nature of science in very dangerous ways.