Chapter 1

Roots of Applied Behavior Analysis

Did you know that . . .

- There may be some validity in your mother's claim that "You're just like your father?"
- Chemicals in your brain may affect your behavior?
- Pretzels preceded M&Ms as rewards for good behavior?
- Benjamin Franklin used applied behavior analysis?

Chapter Outline

Why do people behave as they do? Why do some people behave in socially approved ways and others in a manner condemned or despised by society? Is it possible to predict what people are likely to do? What can be done to change behavior that is harmful to an individual or destructive to society?

In an effort to answer questions like these, human beings have offered explanations ranging from possession by demons to abnormal quantities of chemicals in the brain. Suggested answers have been debated, written about, attacked, and defended for centuries and continue to be offered today. There are good reasons for continuing to investigate human behavior. Information about the development of certain behaviors in human beings may help parents and teachers find the best way of childrearing or teaching. If we know how people are likely to behave under certain conditions, we can decide whether to provide or avoid such conditions. Those of us who are teachers are particularly concerned with changing behavior; that is, in fact, our job. We want to teach our students to do some things and to stop doing others.

To understand, predict, and change human behavior, we must first understand how human behavior works. We must answer as completely as possible the "why" questions asked above. Therefore, Alexander Pope's dictum that "the proper study of mankind is man" (perhaps rephrased to "the proper study of humanity is people") needs no other revision; it is as true in the 21st century as it was in the 18th.

In this chapter we consider the requirements for meaningful and useful explanations of human behavior. We then describe several interpretations of human behavior that have influenced large numbers of practitioners, including teachers. The discussion traces the historical development of a way to understand and predict human behavior called **applied behavior analysis.***

The Usefulness of Explanations

A useful theory has inclusiveness, verifiability, predictive utility, and parsimony.

If a way of explaining behavior is to be useful for the practitioner, it must meet four requirements. First, it should be *inclusive*. It must account for a substantial quantity of behavior. An explanation has limited usefulness if it fails to account for the bulk of human behavior and thus makes prediction and systematic change of behavior impossible. Second, an explanation must be *verifiable*; that is, we should be able to test in some way that it does account for behavior. Third, the explanation should have *predictive utility*. It should provide reliable answers about what people are likely to do under certain circumstances, thereby giving the practitioner the opportunity to change

* Words printed in **boldface** in the text are defined in the glossary at the end of the book.

behavior by changing conditions. Fourth, it should be *parsimonious*. A parsimonious explanation is the simplest one that will account for observed phenomena. Parsimony does not guarantee correctness (Mahoney, 1974) because the simplest explanation may not always be the correct one, but it prevents our being so imaginative as to lose touch with the reality of observed data. When the bathroom light fails to operate at 3 a.m., one should check the bulb before calling the electric company to report a blackout. There may be a blackout, but the parsimonious explanation is a burned-out bulb. In examining some of the theories developed to explain human behavior, we shall evaluate each explanation for its inclusiveness, verifiability, predictive utility, and parsimony.

Biophysical Explanations

Some theorists contend that human behavior is controlled by physical influences.

Since physicians of ancient Greece first proposed that human behavior was the result of interactions among four bodily fluids or "humors"—blood, phlegm, yellow bile (choler), and black bile (melancholy)—theorists have searched for explanations for human behavior within the physical structure of the body. Such theories have included those based on genetic or hereditary factors, those that emphasize biochemical influences, and those that suggest aberrant behavior is caused by some damage to the brain. The following anecdote indicates a belief in hereditary influences on behavior.

Professor Grundy Traces the Cause

Having observed an undergraduate student's behavior for some time, Professor Grundy noticed that the student was consistently late for class (when he came at all), invariably unprepared, and frequently inattentive. Because Grundy was certain his dynamic, meaningful lectures were not related to this behavior, he decided to investigate the matter. He paid a visit to the high school attended by the student and located his 10th-grade English teacher, Ms. Marner. "Yes, DeWayne was just like that in high school," said Ms. Marner. "He just didn't get a good background in middle school."

Professor Grundy then went to visit the middle school. "You know," said the guidance counselor, "a lot of our kids are like that. They just don't get the foundation in elementary school." At the elementary school, Professor Grundy talked to the principal. "DeWayne was like that from day one. His home situation was far from ideal. If we don't have support from the home, it's hard to make much progress."

Professor Grundy, sure that he would at last find the answer, went to talk to DeWayne's mother. "I'll tell you," said DeWayne's mother, "he takes after his father's side of the family. They're all just like that."

Genetic and Hereditary Effects

DeWayne's mother explained his inappropriate behavior by referring to hereditary influences. Could she have been right? The effects of heredity on human behavior, both typical and atypical, have been investigated extensively. There is little question that mental retardation, which results in significant deficits in a wide range of behaviors, is sometimes associated with chromosomal abnormalities or with the inheritance of recessive genes (Patton, Payne, & Beirne-Smith, 1990). Evidence indicates that other behavioral characteristics have some genetic or hereditary basis as well. It is generally accepted that persons with autism have abnormalities in brain development and neurochemistry and that there may be genetic factors related to this disorder (Courchesne, 2004; Freitag, 2007; Prater & Zylstra, 2002). Many emotional and behavior disorders, such as anxiety

disorder, depression schizophrenia, oppositional defiant disorder, and conduct disorder appear to have some genetic origin (Bassarath, 2001; Burke, Loeber, & Birmaher, 2002). Attention deficit disorder and attention deficit hyperactivity disorder also appear to be genetically related (Larsson, Larsson, & Lichtenstein, 2004), as do some learning disabilities (Raskind, 2001).

In addition, inheritance appears to affect some behavioral characteristics that are not necessarily labeled *deviant* or *atypical*. Thomas and Chess (1977) conducted a study of 136 children whose development was closely monitored for a number of years. The authors identified nine categories of behavior that they labeled *temperament*. The categories included activity level, rhythmicity (regularity), approach or withdrawal, adaptability, intensity of reaction, threshold of responsiveness (sensitivity to stimuli), quality of mood (disposition), distractibility, and attention span, and persistence. That these aspects of temperament are observable shortly after birth and remain consistent throughout childhood indicates that they have some constitutional, if not genetic, basis. There is evidence that some clusters of temperamental characteristics may predispose children to be "difficult", but that environmental factors such as childrearing practices have an equal or greater influence on development.

When DeWayne's mother explained her son's behavior to Professor Grundy, her claim that DeWayne takes after his father's family may have involved a degree of truth. It is possible that certain genetic characteristics may increase the probability of certain behavioral characteristics.

Biochemical Explanations

Some children with disabilities show biochemical abnormalities.

Some researchers have suggested that certain behaviors may result from excesses or deficiencies of various substances found in the body. These chemical substances are labeled differently from those hypothesized by the ancient Greeks but are often held responsible for similar disturbances of behavior.

Biochemical abnormalities have been found in some children with serious disturbances of behavior labeled *autism* or *childhood psychosis* (Boullin, Coleman, O'Brien, & Rimland, 1971). Investigation of such factors, however, has established only that biochemical abnormalities exist, not that they cause the disorder.

Other behavior disturbances characterized as hyperactivity, learning disability, or mental retardation have been linked to biophysical factors such as hypoglycemia (Wunderlich, 1977), malnutrition (Cravioto & Delicardie, 1975), and allergic reactions (Feingold, 1975). It is often suggested that biochemical or other physiological factors may, along with other influences, result in damage to the brain or central nervous system.

Professor Grundy Learns to Think in Circles

Professor Grundy, as one of his instructional duties, visited student teachers. On his first trip to evaluate Ms. Harper in a primary resource room, he observed that one student, Ralph, wandered continuously about the room. Curious about such behavior, because the other students remained seated, Professor Grundy inquired, "Why is Ralph wandering around the room? Why doesn't he sit down like the others?" Ms. Harper was aghast at such ignorance on the part of a professor.

"Why, Ralph is hyperactive, Professor Grundy. That's why he never stays in his seat."

"Ah," replied the professor. "That's very interesting. How do you know he's hyperactive?"

With barely concealed disdain, Ms. Harper hissed, "Professor, I know he's hyperactive because he won't stay in his seat."

After observing the class for a few more minutes, he noticed Ms. Harper and the supervising teacher whispering and casting glances in his direction. Professor Grundy once again attracted Ms. Harper's attention. "What," he inquired politely, "causes Ralph's hyperactivity?"

The disdain was no longer concealed. "Professor," answered Ms. Harper, "hyperactivity is caused by brain damage."

"Indeed," responded the professor, "and you know he has brain damage because . . ."

"Of course I know he has brain damage, Professor. He's hyperactive, isn't he?"

Brain Damage

Hyperactivity is not necessarily caused by brain dysfunction.

The circular reasoning illustrated by Ms. Harper is, unfortunately, not uncommon. Many professionals explain a great deal of students' inappropriate behavior similarly. The notion that certain kinds of behavior result from brain damage has its roots in the work of Goldstein (1939), who studied soldiers having head injuries during World War I. He identified certain behavioral characteristics, including distractibility, perceptual confusion, and hyperactivity. Observing similar characteristics in some children with retardation, some professionals concluded that the children must also be brain injured (Strauss & Werner, 1942; Werner & Strauss, 1940) and that the brain injury was the cause of the behavior. This led to the identification of a hyperkinetic behavior syndrome (Strauss & Lehtinen, 1947), assumed to be the result of brain injury. This syndrome included such characteristics as hyperactivity, distractibility, impulsivity, short attention span, emotional lability (changeability), perceptual problems, and clumsiness. Subsequently, the term *minimal brain dysfunction* was used to describe a disorder assumed to exist in children who, although they had no history of brain injury, behaved similarly to those who did. There is, however, little empirical support for using the possibility of brain injury to account for problem behavior in all children who show such behavioral characteristics. Even when brain damage can be unequivocally shown to exist, there is no proof that it causes any particular behavior or that hyperactivity is a result of that damage for any particular individual (Werry, 1986).

Large numbers of children are presently being defined as "at risk" for the development of academic and social problems because of the effects of both influences before birth (such as parental malnutrition or substance abuse) and environmental factors (Davis & McCaul, 1991). In recent years fetal alcohol syndrome (Batshaw & Conlon, 1997), smoking by expectant mothers (Hetherington & Parke, 1986), illegal drug use by expectant mothers (Shriver & Piersal, 1994), and pediatric AIDS (Diamond & Cohen, 1987) have apparently resulted in increased learning and behavioral problems in children. Although there are clear indications that these factors result in biochemical, central nervous system, and other physiological abnormalities, no specific behavioral deficit or excess is directly attributed to any specific factor (Gelfand, Jenson, & Drew, 1988).

The Usefulness of Biophysical and Biochemical Explanations

The search for explanations of human behavior based on physiological factors has important implications. As a result of such research, the technology for preventing or lessening some serious problems has been developed. Perhaps the best-known example of such technology is the routine testing of all infants for phenylketonuria (PKU), an hereditary disorder of metabolism. Placing infants with PKU on special diets can prevent the mental retardation formerly associated with this disorder (Berry, 1969). It is possible that future research may explain a good deal more human behavior on a

biological or hereditary basis. Currently, however, only a small part of the vast quantity of human behavior can be explained in this way.

Some biophysical explanations are testable, meeting the second of our four requirements for usefulness. For example, scientists can definitely establish the existence of Down syndrome by observing chromosomes. Some metabolic or biochemical disorders can also be scientifically verified. Verification of such presumed causes of behavior as minimal brain dysfunction, however, is not dependable (Werry, 1986).

Even with evidence of the existence of some physiological disorder, it does not follow that any specific behavior is automatically a result of the disorder. For the teacher, explanations based on presumed physiological disorders have little predictive utility. To say that Rachel cannot walk, talk, or feed herself because she is developmentally delayed as a result of a chromosomal disorder tells us nothing about the conditions under which Rachel might learn to perform these behaviors. Ms. Harper's explanation of Ralph's failure to sit down on the basis of hyperactivity caused by brain damage does not provide any useful information about what might help Ralph learn to stay in his seat. To say that Harold cannot read because he is a child at risk is to put Harold at the greater risk of not learning because we have low expectations for him. Even apparently constitutional differences in temperament are so vulnerable to environmental influences (Thomas & Birch, 1984) that they provide only limited information about how a child is apt to behave under given conditions.

The final criterion, parsimony, is also frequently ignored when physical causes are postulated for student behaviors. Searching for such causes often distracts teachers from simpler, more immediate factors that may be controlling behaviors in the classroom. Perhaps the greatest danger of such explanations is that some teachers may use them as excuses not to teach: Rachel cannot feed herself because she is developmentally delayed, not because I have not taught her. Ralph will not sit down because he is brain damaged, not because I have poor classroom management skills. Irving cannot read because he has dyslexia, not because I have not figured out a way to teach him. Biophysical explanations may also cause teachers to have low expectations for some students. When this happens, teachers might not even try to teach things students are capable of learning. The accompanying chart summarizes the usefulness of biophysical theory.

The Usefulness of Biophysical Theory

	Good	Fair	Poor
Inclusiveness			✓
Verifiability		✓	
Predictive Utility			✓
Parsimony			✓

Developmental Explanations

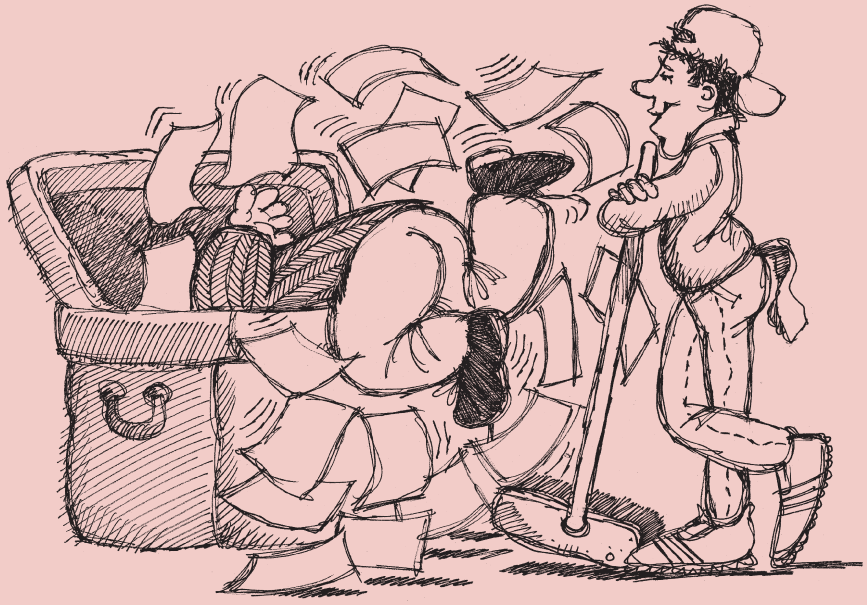
Observation of human beings confirms that many predictable patterns of development occur. Physical growth proceeds in a fairly consistent manner. Most children start walking, talking, and performing some social behaviors such as smiling in fairly predictable sequences and at generally predictable chronological ages (Gesell & Ilg, 1943). Some theorists have attempted to explain many aspects of human behavior—cognitive, social, emotional, and moral—based on fixed, innate developmental sequences. Their proposed explanations are meant to account for normal as well as "deviant" (other than the accepted or usual) human behavior. The following sections review two of the numerous

developmental theories and examine their usefulness in terms of inclusiveness, verifiability, predictive utility, and parsimony.

A Freudian by the Garbage Can

Upon returning to the university after observing student teachers, Professor Grundy prepared to return to work on his textbook manuscript, now at least 7 months behind schedule. To his horror, his carefully organized sources, notes, drafts, and revisions were no longer "arranged" on the floor of his office. Worse, his carefully organized sticky notes had been removed from the walls, door, windows, and computer. Professor Grundy ran frantically down the hall, loudly berating the custodial worker who had taken advantage of his absence to remove what he considered "that trash" from the room so that he could vacuum and dust.

As Grundy pawed through the outside garbage can, a colleague offered sympathy. "That's what happens when an anal-expulsive personality conflicts with an anal-retentive." Grundy's regrettably loud and obscene response to this observation drew the additional comment, "Definite signs of regression to the oral-aggressive stage there, Grundy."



"Well, well, Professor Grundy, did you lose something or are you just doing 'research' on the things you professors throw away?"

Psychoanalytic Theory

Although many different explanations of human behavior have been described as psychoanalytic, all have their roots in theories of Sigmund Freud (Fine, 1973), who described human behavior in an essentially developmental manner (Kessler, 1966). Freud's assertion that normal and aberrant human behavior may be understood and explained on the basis of progression through certain crucial stages (Hall, 1954) is perhaps the most commonly accepted and most widely disseminated of his theories. The hypothetical stages include oral (dependent and aggressive), anal (expulsive and retentive), and phallic (when gender awareness occurs). These stages are believed to occur before

the age of 6 and if mastered, result in emergence into the latency stage, which represents a sort of rest stop until puberty, when the last stage, the genital stage, emerges.

This theory suggests that people who progress through the stages successfully become relatively normal adults. In Freud's view, problems arise when a person fixates (or becomes stuck) at a certain stage or when anxiety causes a regression to a previous stage. People who fixate at or regress to the oral-dependent stage may merely be extremely dependent, or they may seek to solve problems by oral means such as overeating, smoking, or alcohol or drug abuse. A person fixated at the oral-aggressive stage may be sarcastic or verbally abusive. Fixation at the anal-expulsive stage results in messiness and disorganization; at the anal-retentive stage, in compulsive orderliness.

A Stage Theory of Cognitive Development

Jean Piaget was a biologist and psychologist who proposed a stage theory of human development. Piaget's descriptions of the cognitive and moral development of children have had extensive impact among educators. Like Freud, Piaget theorized that certain forces, biologically determined, contribute to development (Piaget & Inhelder, 1969). The forces suggested by Piaget, however, are those enabling the organism to adapt to the environment—specifically, assimilation, the tendency to adapt the environment to enhance personal functioning, and accommodation, the tendency to change behavior to adapt to the environment. The process of maintaining a balance between these two forces is called equilibration. Equilibration facilitates growth; other factors that also do so are organic maturation, experience, and social interaction. Piaget's stages include sensory-motor (birth to 11/2 years), preoperational (11/2 to 7 years), concrete operations (7 to 11 years), and formal operations (12 years to adulthood) (Dember & Jenkins, 1970).

The Usefulness of Developmental Explanations

Both developmental theories we have discussed are inclusive; they apparently explain a great deal of human behavior, cognitive and affective, normal and atypical. Verifiability, however, is another matter. Although Piagetian theorists have repeatedly demonstrated the existence of academic and preacademic behaviors that appear to be age related in many children (Piaget & Inhelder, 1969), attempts to verify psychoanalytic explanations have not been successful (Achenbach & Lewis, 1971). Considerable resistance to verifying theoretical constructs exists among those who accept the psychoanalytic explanation of human behavior (Schultz, 1969). Although it can be verified that many people act in certain ways at certain ages, this does not prove that the cause of such behavior is an underlying developmental stage or that failure to reach or pass such a stage causes inappropriate or maladaptive behavior. There is little evidence to verify that the order of such stages is invariant or that reaching or passing through earlier stages is necessary for functioning at higher levels. (Phillips & Kelly, 1975). The accompanying chart summarizes the usefulness of developmental theory.

The Usefulness of Developmental Theory

	Good	Fair	Poor
Inclusiveness	✓		
Verifiability			✓
Predictive Utility		✓	
Parsimony			✓

Some developmental theories can predict what some human beings will do at certain ages. By their nature these theories offer general information about average persons. However, "a prediction about what the average individual will do is of no value in dealing with a particular individual" (Skinner, 1953, p. 19). Developmental theories do not provide information about what conditions predict an individual's behavior in specific circumstances. The practitioner who wishes to change behavior by changing conditions can expect little help from developmental theories. Developmental explanations of behavior are equally inadequate when judged by the criterion of parsimony. To say that a child has temper tantrums because he is fixated at the oral stage of development is seldom the simplest explanation available. Because of their lack of parsimony, developmental explanations may lead the teacher to excuses as unproductive as those prompted by biophysical explanations. Teachers, particularly teachers of students with disabilities, may wait forever for a student to become developmentally ready for each learning task. An explanation that encourages teachers to take students from their current levels to subsequent levels is clearly more useful than a developmental explanation—at least from a practical point of view. We might expect Professor Grundy's developmental colleagues, for example, to explain Grundy's difficulty with the concept of hyperactivity on the basis of his failure to reach the level of formal operational thinking required to deal with hypothetical constructs. Might there be a more parsimonious, more useful explanation of his behavior? Professor Grundy continues to collect theories of behavior in the following episode.

Professor Grundy Gains Insight

Having been thoroughly demoralized by his interaction with his student teacher, Professor Grundy decided to pay another surprise visit that afternoon. He was determined to avoid subjecting himself to further ridicule. He did not mention Ralph's hyperactivity but instead concentrated on observing Ms. Harper's teaching. Her lesson plan indicated that she was teaching math, but Professor Grundy was confused by the fact that her group was playing with small wooden blocks of various sizes. Ms. Harper sat at the table with the group but did not interact with the students.

At the conclusion of the lesson, Professor Grundy approached Ms. Harper and asked her why she was not teaching basic addition and subtraction facts as she had planned.

"Professor," stated Ms. Harper, "I conducted my lesson exactly as I had planned. The students were using the blocks to gain insight into the relationship among numbers. Perhaps you are not familiar with the constructivist approach, but everyone knows that true insight is vital to the learning process and that it is impossible to teach children; we can only facilitate their own inner construction of knowledge."

Cognitive Explanations

The educational theory espoused (in a somewhat exaggerated form, to be sure) by Ms. Harper is based on an explanation of human behavior and learning that combines elements of developmental theory, especially Piagetian, with a theory first described in Germany in the early part of the 20th century. The first major proponent of this explanation was Max Wertheimer (Hill, 1963), who was interested in people's perception of reality.

Wertheimer suggested it was the relationship among things perceived that was important rather than the things themselves. People, he said, tend to perceive things in an

organized fashion, so that what is seen or heard is different from merely the parts that compose it. He labeled an organized perception of this type a *gestalt*, using a German word for which there is no exact English equivalent but which may be translated as "form," "pattern," or "configuration." English-speaking advocates of this view have retained the word *gestalt*, and we call this explanation Gestalt psychology. Koffka (1935) applied Wertheimer's theories to learning as well as perception. He concluded that learning in human beings is also a process of imposing structure on perceived information. Wertheimer also applied gestalt theory to human problem solving. He studied children's and adults' insights into geometric problems and concluded that meaningful solutions depended on insight and that rote learning—even if it led to correct solutions to problems—was less useful.

Gestalt psychology has had considerable influence on education. The best-known educator to espouse this approach to understanding behavior is Jerome Bruner (1960). What has come to be called the cognitive theory of education places an emphasis on rearranging thought patterns and gaining insight as a basis for learning new academic and social behaviors. The resulting teaching practices are called discovery learning. Learning is explained on the basis of insight, pattern rearrangement, and intuitive leaps. Teachers do not impart knowledge; they merely arrange the environment to facilitate discovery. Motivation is presumed to occur as a result of innate needs that are met when organization is imposed on objects or events in the arrangement. Motivation is thus intrinsic and need not be provided by the teacher. In its latest manifestation, cognitive theory applied to education has been termed *constructivism*. This approach holds that teachers cannot provide knowledge to students; students must construct their own knowledge in their own minds (Brooks, 1990). "Rather than behaviours or skills as the goal of instruction, concept development and deep understanding are the foci" (Fosnot, 1996, p. 10).

Educators who espouse gestalt theory encourage "discovery learning."

Principles derived from Gestalt psychology have also been applied to social behavior, notably in the work of Lewin (1951). His approach has been called *field theory* or *cognitive field theory*. Lewin described human social behavior as based on factors within the person's "life space," the environment as it is perceived by the person and as it affects the person's behavior. He asserted that different people perceive and value environmental objects and events in different ways and that forces exist within people to move them toward or away from these objects or events. Based on a complex procedure for "mapping" or drawing diagrams of people's life spaces, Lewin stated that predictions could be made about what people would do based on the value of the events and the strength of the force. Changing behavior thus depends on changing people's perceptions of their life space and the relationships among the various events and objects in it.

The Usefulness of Cognitive Explanations

Cognitive theory explains a great deal of human behavior. Theorists can account for both intellectual and social behavior. Virtually all behavior can be explained as the result of imposing structure on unstructured environmental events or of perceiving the relative importance of such events. Thus, cognitive theory meets the criterion of inclusiveness.

The theory lacks verifiability, however. Because all the processes that are supposed to take place occur internally, there is no way to confirm their existence. Only the outcome is verifiable—the process is assumed.

The predictive utility of cognitive theory is also extremely limited. In academic areas, the teacher who uses a discovery or constructivist approach has very little control over what students will discover or construct. Most advocates of this approach would insist

that they do not want to predict outcomes of learning. Unfortunately, this unwillingness to control the outcome of the teaching–learning process has led to rather poor results. Educational practices based on a cognitive approach have been less successful than those emphasizing direct instruction (Engelmann & Carnine, 1982).

The predictive utility of cognitive field theory is somewhat greater than that of cognitive theory. If we know enough about the objects and events in a person's life space, the value that she assigns them, and her motivation to approach or avoid them, we may be able to predict behavior. Given all this information, of course, we could almost certainly predict behavior without recourse to theory.

Addressing our final criterion, we must conclude that cognitive theory is not parsimonious. In neither intellectual nor social areas are the explanations necessary to understanding or predicting behavior.

The Usefulness of Cognitive Theory

	Good	Fair	Poor
Inclusiveness	✓		
Verifiability			✓
Predictive Utility			✓
Parsimony			✓

Although all the theories described so far provide information about human behavior, none of them meets all four of our criteria. The explanations we have provided are very general, and our conclusions about their usefulness should not be taken as an indication that they have no value. We simply believe they provide insufficient practical guidance for classroom teachers. After the following vignette, we shall describe a behavioral explanation of human behavior that we believe most nearly reaches the criteria of inclusiveness, verifiability, predictive utility, and parsimony.

Professor Grundy Takes Action

Professor Grundy had an absolutely rotten day. A number of the students in his 8 a.m. class—including, of course, DeWayne—had come in late, disrupting his lecture. He had been ridiculed by a student teacher; his precious manuscript had been retrieved from the dumpster in a sadly wrinkled and malodorous condition; his colleague had made repeated references to "anal-expulsive" and "oral-aggressive" tendencies during the day in spite of Grundy's protests.

After arriving at home and pouring himself a large drink for medicinal purposes, Grundy decided something must be done. He made several detailed plans and retired for the evening, confident he was on the right track. The next morning he arose, enthusiastically determined, in spite of a slight headache, to put his plans into action.

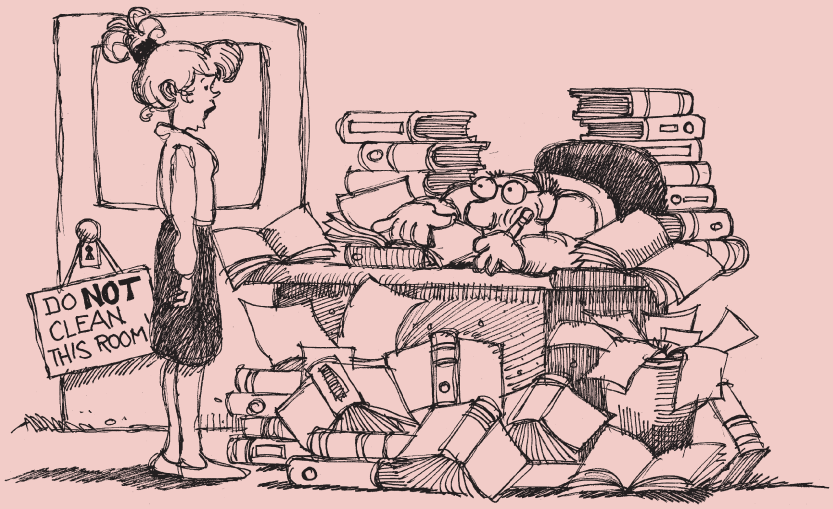
His first step was to arrive at his 8 a.m. class 5 minutes early—somewhat of a novelty because he usually arrived several minutes late. He spent the extra 5 minutes chatting affably with students and clarifying points from the previous day's lecture when asked to do so. At 8:00 sharp, he presented each of the five students present with an "on-time slip" worth 2 points on the next exam.

After the morning lecture, Professor Grundy proceeded to his office, where he affixed to the door a large sign reading "PLEASE DO NOT CLEAN THIS OFFICE TODAY." He then

opened the window, wondering just what the biology department had deposited in the dumpster to cause so strong a smell. He spent an hour reorganizing his notes.

Next, Grundy once again visited Ms. Harper, this time suggesting that she would receive an unsatisfactory grade for student teaching unless she learned to control Ralph's behavior and to teach basic math facts. Her habitual expression of disdain changed to one of rapt attention. Professor Grundy had observed that Ralph, because he was too "hyperactive" to remain in his seat, spent the time while other students worked wandering from toy to toy in the free-time area of the classroom. He suggested that Ms. Harper allow Ralph to play with the toys only after remaining in his seat for a specified length of time: very short periods at first, gradually increasing in length. Grundy further suggested the student teacher make flash cards of basic addition and subtraction facts, allowing the students to play with the colored blocks after they had learned several combinations.

"The librarian sent me over to pick up one of your 48 overdue books . . . if you're . . . uh . . . done with it."



Returning happily to his office, the professor encountered his psychoanalytically oriented colleague, who once again jocularly repeated his insights into Grundy's character. Ignoring the comments, the professor began an animated conversation with the departmental office associate, praising the rapidity with which she was helping him reorganize his manuscript. She assured him it had first priority, because she couldn't wait to be rid of the stinking pages.

Within a short time, Professor Grundy felt that he had things under control. Most of the students enrolled in his 8 a.m. class were present and on time every morning, even though Grundy had begun to give "on-time slips" only occasionally. Ms. Harper had stopped sneering and started teaching. Ralph's wandering had decreased dramatically, and the math group had learned to add and subtract. Grundy continued to ignore his colleague's comments, which gradually ceased when no response was forthcoming, and his notes and drafts were rapidly being transformed into a freshly processed manuscript. The only negative outcome was a sharp note from campus security stating that the condition of his office constituted a fire hazard and that it must be cleaned immediately.

Behavioral Explanations

In the preceding vignette, Professor Grundy emerged as the behaviorist that he is. To solve some of his problems, he used techniques derived from yet another explanation of human behavior. The behavioral explanation states that human behavior, both adaptive and maladaptive, is learned. Learning occurs as a result of the consequences

of behavior. To put it very simply, behavior that is followed by pleasant consequences tends to be repeated and thus learned. Behavior that is followed by unpleasant consequences tends not to be repeated and thus not learned. By assuming that his students, including DeWayne, came to class late, that the custodian cleaned, that the student teacher ridiculed, that Ralph wandered, and that his psychoanalytic colleague teased because they had learned to do so, Professor Grundy was able to teach them to do other things instead. In doing so, he applied several learning principles underlying the behaviorists' view of human behavior. The following sections introduce these principles, each of which will be discussed in detail in later chapters.

Positive Reinforcement

Chapter 7 describes reinforcement in detail.

Positive reinforcement describes a functional relation between two environmental events: a **behavior** (any observable action) and a **consequence** (a result of that action). Positive reinforcement is demonstrated when a behavior is followed by a consequence that increases the behavior's rate of occurrence.

Many human behaviors are learned as a result of positive reinforcement. Parents who praise their children for putting away toys may teach the children to be neat; parents who give their children candy to make them stop screaming in the grocery store may teach the children to scream. The cleaning behavior of Professor Grundy's custodian undoubtedly was learned and maintained through positive reinforcement, as was the wit of Grundy's psychoanalytic colleague. Grundy used positive reinforcement (on-time tickets, conversation, and time with toys) to increase his students' rate of coming to class on time and the amount of time Ralph stayed in his seat.

Negative Reinforcement

Negative reinforcement describes a relationship among events in which the rate of a behavior's occurrence increases when some (usually aversive or unpleasant) environmental condition is removed or reduced in intensity. Human beings learn many behaviors when acting in a certain way results in the termination of unpleasantness. Professor Grundy, for example, learned that opening windows results in the reduction of unpleasant odors in closed rooms. Similarly, the office associate reorganized his manuscript rapidly because when she finished, she could throw away the smelly papers.

Punishment

Chapter 8 describes punishment and extinction in detail.

The word **punishment** also describes a relationship: a behavior is followed by a consequence that decreases the behavior's future rate of occurrence. An event is described as a **punisher** only if the rate of occurrence of the preceding behavior decreases. Behaviorists use the word *punishment* as a technical term to describe a specific relationship; confusion may arise because the same word is used in a nontechnical sense to describe unpleasant things done to people in an effort to change their behavior. To the behaviorist, punishment occurs only when the preceding behavior decreases. In the technical sense of the term, something is not necessarily punishment merely because someone perceives the consequent event as unpleasant. A behaviorist can never say, "I punished him, but it didn't change his behavior," as do many parents and teachers. It is punishment only if the functional relationship can be established. People could say that Professor Grundy's verbal threat to Ms. Harper, for example, was apparently a punisher: her ridiculing comments to him stopped. Of course, we wish he had used a more positive approach.

Extinction

When a previously reinforced behavior is no longer reinforced, its rate of occurrence decreases. This relationship is described as **extinction.** Recall from our vignette that when Grundy no longer reacted to his colleague's ridicule, the behavior stopped. For a behaviorist, all learning principles are defined on the basis of what actually happens, not what we think is happening. Grundy may have thought he was punishing his colleague by yelling or otherwise expressing his annoyance. In reality, the rate of the behavior increased when Grundy reacted in this way; the real relationship was that of positive reinforcement. The behavior stopped when the positive reinforcer was withdrawn.

Antecedent Control

Stimulus control is the focus of Chapter 9.

Requirements that a functional assessment or analysis be performed for students with disabilities before changes in placement can be made (see Chapter 6 for a detailed discussion) have greatly increased interest in antecedent control. Teachers and researchers have come to rely much more frequently on examination of antecedent events and conditions, those occurring before the behavior, to determine what might be setting the stage for appropriate or challenging behaviors. There is also increased emphasis on manipulating antecedent conditions or events to manage behavior.

An antecedent that occurs immediately before a behavior is called a discriminative stimulus and is said to "occasion" (to set the occasion for) a behavior. There is a functional relation, called **stimulus control,** between behavior and an **antecedent stimulus** rather than behavior and its consequences. Consequences must have been present during the development of the relation, but the antecedent condition or event now serves as a signal or cue for the behavior. In our vignette, the custodian's adherence to posted notices had apparently been reinforced in the past, so Professor Grundy's sign was effective even in the absence of a reinforcer or a punisher.

Recently, researchers have been investigating more distant varieties of antecedent events and conditions (Smith & Iwata, 1997). Often referred to as **setting events** or *establishing operations,** these conditions or events may occur simultaneously with a discriminative stimulus or hours or even days before (Horner, Vaughn, Day, & Ard, 1996). They may occur in the same setting or in a completely different one. They influence behavior by temporarily changing the value or effectiveness of reinforcers. The simplest kinds of setting events to describe are deprivation and satiation. A student who has just come in from the playground, sweating buckets from playing a hard game of kickball, is likely to be more responsive to a soft drink as a potential reinforcer than one who has just consumed a soda in the air-conditioned cafeteria. Setting events, however, can be much more complex. Kazdin (2000) described three types of setting events: social, physiological, and environmental. Bailey, Wolery, and Sugai (1988) subdivided environmental setting events into instructional dimensions, physical dimensions, social dimensions, and environmental changes. These varieties of conditions and events may include variables as diverse as a noisy or uncomfortably warm classroom (environmental), the presence of a disliked staff member or peer (social), or a headache (physiological).

Bailey et al. (1988) included considerations about instructional materials that may not be age appropriate or gender appropriate. It may be that no reinforcer will (or indeed

* Some authors (Michael, 2000) use the term *establishing operations*; others (Kazdin, 2000) use *setting events*. Some use the terms interchangeably; others differentiate between them. Kennedy and Meyer (1998) suggest that the terms describe similar concepts from different theoretical perspectives. Until the confusion about terminology (sometimes within a single edited text) is cleared up (Horner & Harvey, 2000), we prefer to use the term *setting events*.

should) induce a teenage boy to touch, much less read, a colorfully illustrated book about the little mermaid. We believe that issues of students' ethnic or cultural heritage can also serve as setting events. Students are much more motivated to interact with materials that portray people like themselves (Gay, 2002). Attention to cultural diversity may enable teachers to provide reinforcers that are more meaningful and powerful and to avoid strategies that are ineffective or offensive. Strategies such as personalized contextual instruction (Voltz, 2003) that embed instruction into contexts of interest to students in a given setting (Voltz, 1999) may enhance the value of reinforcers. The following anecdote describes a classroom using this approach.

Music Hath Charms

Ms. Garcia, a general education teacher, Mr. Walden, a special education teacher, and Ms. Nguyen, a paraprofessional, share the responsibility for an inclusive primary class of 25 students that they privately agree "gives new meaning to the term diversity." Their students range in age from 7 to 9. They have 14 boys and 11 girls: 12 African-American students, 8 Hispanic students, and 4 Asian students. They have 7 children with learning disabilities, 4 children with behavior disorders, and 2 children who are intellectually gifted. And they have Yuri, a boy from Russia who has autism. What the children have in common is eligibility for free or reduced-price lunch and the fact that all of their teachers believe every one of them is capable of great things.

Things had been going well; the teachers used the standard curriculum and a combination of group and individual teaching. They used a simple point system with the class as a whole (the students could earn tangibles and activities for completing work and behaving appropriately) and implemented more complex positive behavior support plans with some children with more challenging behaviors. The students were making good academic progress but, as Mr. Walden stated at a meeting one afternoon, "Nobody seems real excited about school except us." The three teachers decided to implement an integrated unit approach that Ms. Garcia had learned about in a class she was taking at the local university and researched on the Internet and at the university library. The next morning Ms. Garcia explained the plan to the students, asking them to think about what they would like to study. The students seemed to think the teachers must be kidding and made several suggestions ranging from sports to dinosaurs, but most of the interest appeared to center around music. "Rap!" shouted several students. "Salsa!" suggested others. "All right," agreed Ms. Garcia, as Ms. Nguyen and Mr. Walden moved around praising students who were attending, "Let's make a list of what we already know about music and then a list of things we would like to know. Ms. Nguyen, would you help Yuri put the sticky notes with our ideas on the board?"

After almost an hour they had a good list to start out with and the teachers were startled to see that it was almost lunchtime. They were even more startled to realize that no one had given the students points all morning and that verbal praise and pats on the back had been enough.

Other Learning Principles

In addition to these major learning principles, Professor Grundy illustrated the use of several other influences on human behavior described by behaviorists. These influences include **modeling** and **shaping.** Modeling is the demonstration of behavior. The professor had been modeling inappropriate behavior—coming to class late—and his students had apparently been imitating that behavior. People learn many behaviors, both appropriate and inappropriate, by imitating a model. Infants learn to talk by imitating their caregivers; adults can learn to operate complex machinery by watching a demonstration. Bandura (1977) developed a social learning theory that describes the power of observation in facilitating learning.

Shaping uses the reinforcement of successive approximations to a desired behavior to teach new behavior. Grundy suggested that Ms. Harper use shaping to teach Ralph to stay in his seat. She was initially to reinforce sitting behavior when it occurred for short periods of time and gradually increase the sitting time required for Ralph to earn the reinforcer.

Many behaviors are taught by shaping. Parents may praise a young child effusively the first time she dresses herself, even if her blouse is on inside out and her shorts are on backward. Later she may earn a compliment only if her outfit is perfectly coordinated.

The Task of the Behaviorist

If you can see it, hear it, feel it, or smell it, it's observable. If you can count it or measure it, it's quantifiable.

Behaviorists explain the development of both typical and atypical human behavior in terms of the principles just described. An important aspect of this approach is its emphasis on behavior. To qualify as a behavior, something must be *observable* and *quantifiable* (Baer, Wolf, & Risley, 1968). We must be able to see (or sometimes hear, feel, or even smell) the behavior. To make such direct observation meaningful, some way of measuring the behavior in quantitative terms (How much? How long? How often?) must be established. Behaviorists cannot reliably state that any of the relations described as learning principles exist unless these criteria are met.

Skinner (1953) suggested that behaviorists are less concerned with explaining behavior than with describing it. The emphasis, he states, is on which environmental factors increase, decrease, or maintain the rate of occurrence of specific behaviors. It is important to note that behaviorists do not deny the existence of physiological problems that may contribute to some behavioral problems. Nor do most behaviorists deny the effects of heredity (Mahoney, 1974) or even developmental stages (Ferster, Culbertson, & Boren, 1975). Their primary emphasis, however, is on present environmental conditions maintaining behavior and on establishing and verifying functional relations between such conditions and behavior.

The Usefulness of Behavioral Explanations

One of the most common criticisms of the behavioral approach is that it leaves much of human behavior unexplained. Emphasis on observable behavior has led many to assume that behavioral principles cannot account for any but simple motor responses. However, Skinner (1953, 1957, 1971) applied basic learning principles to explain a wide variety of complex human behavior, including verbal behavior and sociological, economic, political, and religious beliefs.

The fact that behavioral principles have not accounted for all aspects of human behavior should not lead to the assumption that they cannot. In the years since Skinner first identified the principles of behavior that developed into the discipline of applied behavior analysis, many aspects have been accounted for. Many phenomena have yet to be explained. "In the meantime—which may last forever—the best strategy is to isolate variables that influence important behavior and manipulate those variables to make life better" (Poling & Byrne, 1996, p. 79). Because behaviorists refuse to theorize about what they have not observed, explanation must await verification. Behaviorists are ready temporarily to sacrifice some degree of inclusiveness for verifiability.

Verifiability is the essence of the behavioral explanation. Other theorists posit a theory and attempt to verify it through experimental investigation. Behaviorists, on the other hand, investigate before formulating what may be described as generalizations rather than theories. That adult attention serves as a positive reinforcer for most children (Baer & Wolf, 1968; Harris, Johnston, Kelley, & Wolf, 1964) is an example of such a generalization. This statement was made only after repeated observations established a functional relation between children's behavior and adult attention. The accompanying chart summarizes the usefulness of behavioral theory.

The focus of the behavioral approach is changing behavior. Predictive utility is an essential part of any behavioral explanation. Functional relations are established and

The Usefulness of Behavioral Theory

	Good	Fair	Poor
Inclusiveness		✓	
Verifiability	✓		
Predictive Utility	✓		
Parsimony	✓		

generalizations are made precisely so that they can be used to change maladaptive or inappropriate behavior and increase appropriate behavior. Behaviorists are reinforced by changing behavior, not by discussing it. Unless it is possible to use generalizations to predict what people will do under certain conditions, behaviorists see little point in making the statements. An enormous body of evidence exists, representing the application of learning principles to human behavior. Such data make possible the prediction of behavior under a wide variety of conditions.

"Explanatory fictions" explain nothing. Behaviorists explain behavior on the basis of observation, not imagination.

Behavioral explanations are parsimonious, satisfying our fourth criterion for usefulness. Describing behavior solely in terms of observable, verifiable, functional relations avoids the use of "explanatory fictions." Such fictions are defined only in terms of their effects, resulting in the circular reasoning we discussed earlier. Rather than invoking "hyperactivity"—an example of an explanatory fiction—to explain Ralph's out-of-seat behavior, Professor Grundy chose a behavioral approach to look at what happened before and after Ralph left his seat. In this way, behaviorism avoids explanations distant from observed behavior and its relationship to the environment. It is unacceptable to explain out-of-seat behavior by labeling the cause as hyperactivity or to explain messiness as fixation at or regression to the anal-expulsive stage of behavior. Neither explanation adds anything useful to our information about the problem.

Haughton and Ayllon (1965) offered one example of the fluency with which many professionals are willing to invoke unparsimonious explanations of behavior. The authors were working with a hospitalized mental patient whose behavior for many years had been limited to sitting and smoking cigarettes. After a period during which smoking was limited, the patient was given cigarettes only when standing up and holding a broom. The patient began carrying the broom most of the time. Two psychiatrists were asked to observe and evaluate the patient's behavior. Both offered lengthy and complex explanations, suggesting that the broom served a function similar to that of a young child's "blankie" or that it represented an infant she wished she had, or the scepter of an omnipotent queen. When staff members stopped giving cigarettes to the woman while she was carrying the broom, she stopped carrying the broom. Although we stated earlier that the parsimonious explanation may not always be correct, in this case it was. Even when the development of unusual behavior is not as easy to trace as in this example, the assumption that such behaviors are being maintained by current environmental conditions and that the behavior may be changed by changing the environment is not merely parsimonious, it is supremely optimistic. The teacher who concentrates on discovering and changing the environmental conditions maintaining students' inappropriate or maladaptive behavior does not give up on them because they have cultural differences, retardation, brain damage, emotional disturbance, hyperactivity, or are at risk, or developmentally unready to learn; she teaches them. If students' behavior is described in terms of behavioral excesses (too much moving around) or deficits (too little reading), as suggested by Gelfand and Hartmann (1975) and Hersen and Bellack (1977), rather than

in terms of explanatory fictions, the teacher can go about the business of teaching—decreasing behavioral excesses and overcoming behavioral deficits.

Historical Development of Behaviorism

Behaviorism as a science has roots in philosophical and psychological traditions originating several centuries ago. The learning principles described earlier certainly existed before being formally defined. People's behavior has been influenced since the beginning of civilization. In the following section, we will examine several historical descriptions of how people have used the relation between behavior and its consequences. Then we will trace the development of behaviorism as a formal way of explaining, predicting, and changing human behavior.

Historical Precedents

The arrangement of environmental conditions in order to influence behavior is by no means a recent invention. It is said that the ancient Romans put eels in the bottom of wine cups to decrease excessive drinking. Crossman (1975, p.348) provided an historical example of the use of positive reinforcement.

> There is a fascinating history behind the pretzel. About 610 an imaginative Alpine monk formed the ends of dough, left over from the baking of bread, into baked strips folded into a looped twist so as to represent the folded arms of children in prayer. The tasty treat was offered to the children as they learned their prayers and thereby came to be called "pretiola"—Latin for "little reward." [From the back of a Country Club Foods pretzel bag, Salt Lake City.]

Educators used behavioral principles long before the principles were formally identified.

Several innovative educators developed elaborate programs of reward and punishment to manage their students' behavior. In the early 19th century, Lancaster (Kaestle, 1973) instituted a system in Great Britain that was later also used in the United States. Students earned tickets that could be exchanged for prizes or money. They lost tickets when they misbehaved.

Benjamin Franklin demonstrated that adults' behavior could also be changed, using a rather different positive reinforcer (Skinner, 1969). When a ship's chaplain complained that few sailors attended prayers, Franklin suggested that the chaplain take charge of serving the sailors' daily ration of rum and deal it out only after the prayers. Attendance improved remarkably.

Behavioral principles operate whether anyone is consciously using them.

Parents and teachers have likewise applied the principles of learning in their efforts to teach children. "Clean up your plate and then you can have dessert," says the parent hoping for positive reinforcement. "When you finish your arithmetic, you may play a game," promises the teacher. Parents and teachers, whether they are aware of it or not, also use punishment: the child who runs into the street is spanked; the student who finishes his assignment quickly is given more work to do. All of us have heard "Just ignore him and he'll stop. He's only doing it for attention." If he does stop, we have an example of extinction. Of course, many parents and teachers extinguish appropriate behavior as well, paying no attention to children who are behaving nicely. Negative reinforcement is demonstrated in many homes every day: "You don't play outside until that room is clean." Teachers also use negative reinforcement when they require students, for example, to finish assignments before going to lunch or to recess. Kindergarten teachers who ask their charges to use their "inside voices" are trying to establish stimulus control. Whenever teachers show their students how to do something, they are modeling.

It becomes apparent that a person does not need to know the names of the relationships involved to use them. Indeed, applying behavioral learning principles sounds a lot

like common sense. If it is so simple, why must students take courses and read books? Why have such quantities of material been written and so much research conducted?

The answer is that it is inefficient to fail to arrange environmental conditions so that functional relations are established, or to allow such relations to be randomly established, or to assume that such relations have been established based only on common sense. This inefficiency has resulted in high levels of maladaptive behavior in schools and sometimes frighteningly low levels of academic and preacademic learning. It is our aim in writing this book to help teachers become applied behavior analysts. The derivation and definition of the term *applied behavior analysis* will be discussed in the remaining sections of this chapter.

Philosophical and Psychological Antecedents

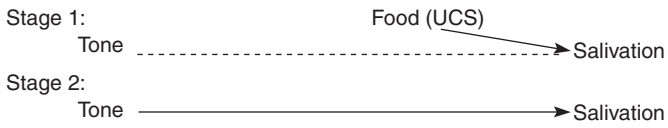
The roots of the behavioral viewpoint are firmly planted in a 19th-century philosophical movement known as *positivism*, which in its turn evolved from the 17th-century writings of Francis Bacon (Smith, 1992). Positivism's earliest proponent, Auguste Comte, emphasized that the only valid knowledge was that which was objectively observable.

A second important contribution came from animal psychology, influenced by the work of Charles Darwin (Boring, 1950), which emphasized the continuity between animal and human behavior and thus suggested that something about human beings could be learned through the careful observation of other animals. Animal psychology focused on the adaptation of physical structures in the body to the environment. This focus led to consideration of mental processes in the same light and to a psychological movement known as *functionalism*.

Functionalism was a third important influence on the development of a behavioral approach to explaining human behavior. William James, whose work was a precursor of behaviorism (Boring, 1950), noted that Dewey and James Angell were also influential in turning the emphasis in American psychology from an introspective, theorizing model to one emphasizing a practical, observational approach.

Respondent Conditioning

Most people are aware of the work of Ivan Pavlov, who observed that when a tone was sounded as dogs were fed, the dogs began to salivate when they heard the tone even when food was not present. (Anyone who feeds dogs can observe a similar phenomenon when the dogs arrive drooling when they hear the food pans being taken from the dishwasher.) Pavlov's precise observation and measurement have served as a model for experimental research to this day. His classic experiment involved pairing food powder (which elicits salivation, an automatic reflex) with a tone that would normally have no effect on dogs' salivation. The presentation of the tone preceded the presentation of the food powder; after repeated pairings, salivation occurred when only the tone was presented (Hill, 1970). The food powder was labeled the *unconditioned stimulus* (UCS); the tone, the *conditioned stimulus* (CS). Salivation is an unconditioned response to food powder and a conditioned response to the tone. The relationship may be represented as shown in the accompanying diagram. The process of pairing stimuli so that an unconditioned stimulus elicits a response is known as *Pavlovian, classical*, or *respondent conditioning*.



Associationism

Another influential experimenter whose research paralleled that of Pavlov was Edward Thorndike. Thorndike studied cats rather than dogs, and his primary interest was discovering associations between situations and responses (Thorndike, 1931). He formulated two laws that profoundly influenced the subsequent development of behavioral science. The Law of Effect (Thorndike, 1905) states that "any act which in a given situation produces satisfaction becomes associated with that situation, so that when the situation recurs the act is more likely than before to recur also" (p. 203). Second is the Law of Exercise, which states that a response made in a particular situation becomes associated with the situation. The relationship of the Law of Effect with the principle of positive reinforcement is obvious. The Law of Exercise is similarly related to the stimulus control principle discussed earlier.

Behaviorism

If we were all Watsonians, we couldn't say, "She hurt my feelings," My mind wandered," or "Use your imagination."

The term *behaviorism* was first used by John Watson (1914, 1919, 1925). Watson advocated the complete abolition of any datum in psychology that did not result from direct observation. He considered such concepts as mind, instinct, thought, and emotion both useless and superfluous. He denied the existence of instinct in human beings and reduced thought to subvocal speech, emotion to bodily responses. A Watsonian behaviorist of our acquaintance once responded to a question by saying, "I've changed my mind (you should excuse the expression)." The true Watsonian does not acknowledge the existence of any such entity as "mind."

Watson and Raynor (1920) conditioned a startle response in a baby, Albert, by pairing a white rat (CS) with a loud noise (UCS). Watson contended that all "emotional" responses such as fear were conditioned in similar ways. In an interestingly related procedure, Jones (1924) desensitized a 3-year-old child who showed a fear response to white rabbits and other white furry objects by pairing the child's favorite foods with the rabbit. This procedure was unfortunately not carried out with Albert, who moved away before his conditioned fear could be eliminated. Albert may still be scared of white rats, which may have created a number of problems in his life, including preventing his employment as a behavioral psychologist. Watson later suggested that Albert might eventually seek Freudian therapy to overcome his strange fears and that his problems might be attributed to an unresolved Oedipal complex (Pierce & Eppling, 1999).

Operant Conditioning

Operant behaviors are emitted voluntarily; respondent behaviors are elicited by stimuli.

The learning principles described at the beginning of this section are those suggested by proponents of an *operant conditioning* model for explaining, predicting, and changing human behavior. The best known operant conditioner was B. F. Skinner (1904–1988), who first distinguished operant from respondent conditioning.

Respondent conditioning, you will recall, deals with behaviors elicited by stimuli that precede them. Most such behaviors are reflexive; that is, they are not under voluntary control. Operant conditioning (sometimes called *instrumental conditioning*), on the other hand, deals with behaviors usually considered voluntary rather than reflexive. Operant conditioners are concerned primarily with the consequences of behavior and the establishment of functional relations between behavior and consequences. The behavioral view described earlier is that of operant conditioning, which will be the emphasis of the entire text.

Skinner's early work was with animals, primarily white rats. In this, he followed in the tradition of earlier behaviorists, to whom this particular animal was so important that one researcher (Tolman, 1932) dedicated a major book to *Mus norvegious albinius*, a strain of white rats. Bertrand Russell, the philosopher, is said to have suggested

facetiously that the different emphases in European (primarily gestalt, introspective, and theorizing) and American (primarily behavioral, active, observational) studies may have resulted from differences in the breeds of rats available. Whereas European rats sat around quietly waiting for insight, American rats were active go-getters, scurrying around their cages and providing lots of behaviors for psychologists to observe.

Skinner also worked with pigeons. He explained (1963) that, while in the military during World War II, he was assigned to a building whose windowsills were frequented by these birds. Because there was very little to do, he and his colleagues began to train the pigeons to perform various behaviors. This subsequently developed into a rather elaborate, successful, although ultimately abandoned before fully operational, project to train pigeons to deliver guided missiles to enemy vessels. The pigeons, of course, were limited to one delivery. Although "Project Pigeon" was a source of personal and professional frustration to Skinner, it is credited with moving his interest firmly and finally from the laboratory into applied settings (Capshew, 1993).

Early application of operant conditioning techniques to human beings was directed toward establishing that the principles governing animal behavior also govern human behavior. The use of these principles to change human behavior—usually called *behavior modification*—did not really emerge in nonlaboratory settings until the 1960s. One of the authors remembers being told in an experimental psychology course in 1961 that there was some indication operant conditioning could be applied to simple human behavior. As an example, the instructor laughingly described college students' conditioning their professor to lecture from one side of the room simply by looking interested only when he stood on that side. The instructor insisted that it would not be possible to modify his behavior in this way, because he was aware of the technique. He was wrong; he was backed into one corner of the room by the end of the next lecture.

At that time, however—in spite of Skinner's (1953) theoretical application of operant conditioning techniques to complex human behavior and pioneer studies such as those of Ayllon and Michael (1959) and Birnbrauer, Bijou, Wolf, and Kidder (1965)—few people anticipated the enormous impact that the use of such principles would have on American psychology and education and on other disciplines, including economics (Kagel & Winkler, 1972). The application of behavior modification in real-life settings had become so prevalent by 1968 that a new journal, the *Journal of Applied Behavior Analysis*, was founded to publish the results of research. In Volume 1, Number 1, of the journal, Baer, Wolf, and Risley (1968) defined applied behavior analysis as the "process of applying sometimes tentative principles of behavior to the improvement of specific behaviors, and simultaneously evaluating whether or not any changes noted are indeed attributed to the process of application" (p. 91).

Applied behavior analysis must deal with socially important, observable behaviors. Relationships between behaviors and interventions must be verified.

Baer and his colleagues (1968) suggested that for research to qualify as applied behavior analysis, it must change socially important behavior, chosen because it needs change, not because its study is convenient to the researcher. It must deal with observable and quantifiable behavior, objectively defined or defined in terms of examples, and clear evidence of a functional relation between the behavior to be changed and the experimenter's intervention must exist. In a more recent retrospective analysis of the progress of applied behavior analysis since 1968, the same authors (Baer, Wolf, & Risley, 1987) suggested that in spite of considerable opposition and in light of many failures of the procedures in real settings, applied behavior analysts should persevere. They stated, "current theory has worked far too well to be abandoned in the face of what are more parsimoniously seen as technological rather than theoretical failures" (p. 325). In other words, we still cannot always make what we know ought to work actually work, but that is a problem of implementation not an indication of the inadequacy of applied behavior analysis as a discipline. Johnston (1996) suggested that a greater

separation of applied research from service delivery might provide more controlled conditions for research and thus enable more progress.

Applied behavior analysis is more rigorously defined than behavior modification. In our earlier vignette, Professor Grundy apparently succeeded in modifying behavior, but he failed to meet the criterion of analysis—he had no way of knowing for sure whether his techniques changed behavior or whether the change was mere coincidence. Keeping data about behavior change (or the lack of it) is a fundamental tenet of applied behavior analysis. It is required for many procedures necessary for students with disabilities, including functional analysis of behavior, discussed in Chapter 6, and Responsiveness to Intervention (RTI) (Bradley, Danielson, & Doolittle, 2007) used as part of the identification for students with learning disabilities. This book is designed to help teachers become applied behavior analysts, effective modifiers of behavior, and efficient analyzers of the principles of learning involved in all aspects of their students' performance.

Teachers who learn and practice the principles of applied behavior analysis can help their students master functional and academic skills in a systematic and efficient manner and can document their students' progress for parents and other professionals. They can manage behavior positively so that their focus remains on learning. They can teach students to get along with peers and adults and to make good choices. By providing learning environments that are safe, joyful, and successful, they can make enormous differences in students' lives.

Summary

We described a number of approaches to explaining human behavior. We evaluated these approaches in terms of their inclusiveness, verifiability, predictive utility, and parsimony. We also described an explanation of human behavior that appears to us to be the most useful—the behavioral explanation.

In tracing the history of the behavioral approach to human behavior, we emphasized the development of a science of applied behavior analysis. We discussed the necessity for concentrating on socially useful studies of human behavior and on careful observation of the establishment of functional relations. We also provided a rationale for learning and using the principles of applied behavior analysis and some examples of their use in various educational settings.

Key Terms

applied behavior analysis
positive reinforcement
behavior
consequence
negative reinforcement
punishment
punisher
extinction
stimulus control
antecedent stimulus
setting events
modeling
shaping

Discussion Questions

1. Why do applied behavior analysts consider biophysical, developmental, and cognitive explanations of behavior less useful than a behavioral explanation?
2. What useful information might teachers gain from biophysical, developmental, or cognitive information about their students?
3. What differentiates the historical examples of the use of consequences to change behavior from applied behavior analysis?
4. What is the difference between operant and respondent conditioning? Why, in a text for teachers, have we chosen to emphasize operant conditioning?

Chapter 2

Preparing Behavioral Objectives

Did you know that . . .

- If a composer orchestrates without the melody in mind, the Dixie Chicks could sound like Santana?
- There are reasons for writing behavioral objectives besides satisfying legal or administrative requirements?
- Aggression is in the eye of the beholder?
- Even professors write behavioral objectives?
- Ninety percent may not be a passing grade?
- Accuracy is not always enough?

Chapter Outline

Definition: A behavioral objective is a statement that communicates a proposed change in behavior. It describes a level of performance and serves as a basis for evaluation.

In this chapter we will discuss the first step in carrying out a program for behavior change: defining the target behavior—the behavior to be changed. A target behavior may be selected because it addresses a behavioral deficit (such as too few math skills) or a behavioral excess (such as too much screaming). After the behavior to be changed has been identified, a written **behavioral objective** is prepared. A behavioral objective describes the behavior that should result from the instruction or intervention that is planned. It describes the intended outcomes of instruction, not the procedures for accomplishing those outcomes (Mager, 1997).

A behavioral objective for a student who demonstrates a deficit in math skills would describe the level of math performance the student should reach. A behavioral objective for a student who screams excessively would describe an acceptable level of screaming. Anyone reading a behavioral objective should be able to understand exactly what a student is working to accomplish. Because behavioral objectives are such an integral part of planning for student behavior change, they are required as part of the individualized education program (IEP) for students with disabilities. We will also talk about the relationship between objectives and the IEP.

You will meet some teachers who are learning to use a behavioral approach in their teaching. Through them, you will encounter some of the difficulties of putting behavioral programs into effect. Consider the plight of Ms. Samuels, the resource teacher, in the following vignette.

Are We Both Talking About the Same Thing?

Ms. Wilberforce, the third-grade teacher, was in a snit.

"That special ed consulting teacher," she complained to her friend, Ms. Folden, "is absolutely useless. I asked her 2 months ago to work on vowels with Martin and he still doesn't know the short sounds."

"You're absolutely right," agreed Ms. Folden, "I told her last September that Melissa Sue had a bad attitude. The longer the special ed teacher sees Melissa Sue, the worse it gets. All Melissa Sue does now is giggle when I correct her. It seems to me that we were better off without special ed teachers."

Meanwhile, Ms. Samuels, the special ed teacher, was complaining bitterly to her supervisor.

"Those general education teachers are so ungrateful. Just look at what I've done with Martin. He can name all the vowels when I ask him, and he even knows a little song about them. And Melissa Sue, who used to pout all the time, smiles and laughs so much now. I've done exactly what the teachers asked—why don't they appreciate it?"

Definition and Purpose

Behavioral objectives improve communication.

The preceding vignette illustrates one of the most important reasons for writing behavioral objectives: to clarify the goals of a student's behavior-change program and thus to facilitate communication among people involved in the program. Because it is a written statement targeting a specific change in behavior, the objective serves as an

agreement among school personnel, parents, and students about the academic or social learning for which school personnel are taking responsibility.

An objective may also serve to inform students of what is expected. It is a statement of proposed student achievement and tells students what they will be learning or in what manner and to what degree their behavior is to change. Providing students with a statement of proposed learning outcomes enables them to match their performance with a standard of correct or expected performance. This allows for ongoing evaluation and provides informative feedback and reinforcement (Gagne, 1985, p. 309).

A second reason for writing behavioral objectives is that a clearly stated target for instruction facilitates effective programming by the teacher and ancillary personnel. A clearly stated instructional target provides a basis for selecting appropriate materials and instructional strategies. Mager (1997) pointed out that "machinists and surgeons do not select tools until they know what they're intending to accomplish. Composers don't orchestrate scores until they know what effects they are trying to create" (p. 14). Clearly written behavioral objectives should prevent the classroom teacher from using materials simply because they are available or strategies simply because they are familiar. The selection of materials and teaching strategies is more likely to be appropriate if objectives are clearly defined.

There is yet another excellent reason for writing behavioral objectives. Consider the following account.

A Matter of Opinion

Mr. Henderson, the teacher of a preschool class for students with developmental delays, hurried to the principal's office in a state of complete panic. The parents of Alvin, one of his students, had just threatened to remove the boy from school. They insisted Mr. Henderson was not teaching Alvin anything and was not making it possible for Alvin to spend more time with his general education kindergarten class. Mr. Henderson had agreed in August to work on toilet training with Alvin and felt the boy had made excellent progress. Alvin's parents, however, were upset because Alvin still had several accidents every week. They and the kindergarten teacher insisted that Mr. Henderson had not reached his stated goal.

"I have toilet trained Alvin," howled Mr. Henderson. "Wouldn't you consider only two or three accidents a week all right?"

Behavioral objectives help evaluate progress.

Mr. Henderson's panic could have been prevented if a clearly written objective statement had come out of the August meeting. If a definition of *toilet trained* had been established at the beginning of the year, there would have been no question as to whether that objective had been achieved. Behavioral objectives provide for precise evaluation of instruction. When a teacher identifies a deficit or an excess in a student's behavioral repertoire, he has identified a discrepancy between current and expected levels of functioning. If the teacher states a performance criterion (the ultimate goal) and records ongoing progress toward this goal, both formative (ongoing) and summative (final) evaluation of intervention procedures become possible so that programs can be changed as necessary and plans made for the future. Ongoing evaluation and measurement enable the teacher, the student, or a third party to monitor progress continuously and to determine when goals have been reached. Continuous monitoring minimizes individual interpretations or prejudices in judgment when instructional procedures or a student's performances are evaluated.

Pinpointing Behavior

Before an objective can be written or a behavior change program initiated, the target behavior must be described clearly. Referral information may often be vague and

imprecise. To write effective objectives, the applied behavior analyst must refine broad generalizations into specific, observable, measurable behaviors. This process is frequently referred to as **pinpointing** behavior.

Pinpointing may be accomplished by asking a series of questions, usually including "Could you please tell me what he *does?"* or "What exactly do you want her to *do?"* For example, teachers often refer students to a behavior analyst because of "hyperactivity." The referring teacher and the applied behavior analyst must define this hyperactive behavior by describing exactly what is occurring. Is it that the student, like Ralph in Ms. Harper's class, wanders around the room? Does he tap his pencil on the desk? Does he weave back and forth in his chair?

Many categories of behavior may result in referrals and require pinpointing. Here are a few examples, with some questions that may help to refine the definition:

Sebastian can't do math: Is the problem that he does not have basic arithmetic computation skills, or that he cannot finish his problems within the time limit set, or that he refuses to attempt the problems?

Stella is always off task: Is the problem that she stares out the window, or that she talks to her neighbor, or that she scribbles in her book instead of looking at the chalkboard?

Chance is always disturbing others: Is he grabbing objects from someone, or talking to others during lessons, or hitting his neighbor, or knocking neighbors' books off the desk, or pulling someone's hair?

Maggie's lab projects are a mess: Is it that she cannot read the instructions in the lab manual, or that her handwriting is sloppy, or that she does not do the prescribed steps in the right order, or that she can do the experiments but cannot write the results coherently?

Laura throws tantrums: Is she crying and sobbing, or is she throwing herself on the floor, or is she throwing objects around the room?

A teacher may ask a similar series of questions in describing more complex or abstract categories of behaviors. If the referring teacher said, "Carol doesn't use critical thinking skills," the applied behavior analyst would want to know if Carol

1. distinguishes between facts and opinions;
2. distinguishes between facts and inferences;
3. identifies cause–effect relationships;
4. identifies errors in reasoning;
5. distinguishes between relevant and irrelevant arguments;
6. distinguishes between warranted and unwarranted generalizations;
7. formulates valid conclusions from written material;
8. specifies assumptions needed to make conclusions true. (Gronlund, 1985, p. 14)

The behavior analyst may need to address other issues. If, for example, a student is out of his chair at inappropriate times, the teacher's concern may be either the number of times he gets out or the length of time he stays out. The student who gets up only once, but stays up all morning, is doing something quite different from the student who hops in and out of his seat every few minutes. Different intervention strategies and data collection techniques are needed. For complex behaviors such as temper tantrums, during which many discrete behaviors may occur simultaneously, it may be helpful to list

the behaviors in some order of priority. They might be listed, for example, in order from least to most interference to the child or to the environment. After referral information has been refined so that target behaviors can be clearly described, **educational goals** and eventually behavioral objectives can be written.

EDUCATIONAL GOALS

Goals precede objectives.

Objectives should be derived from a set of educational goals that provide the framework for the academic year. These goals should evolve from an accumulation of evaluation information and should be correlated with curriculum planning. Goals define the anticipated academic and social development for which the school will take responsibility. During goal selection, educators estimate what proportion of the student's educational potential is to be developed within the next academic year. Thus, educational goals (long-term objectives) are statements of annual program intent, whereas behavior objectives (short-term or instructional objectives) are statements of actual instructional intent, usually for a 3- to 4-month period (quarterly) for individuals with more severe disabilities and for the length of time of the school's grading period for students with mild disabilities.

ESTABLISHING GOALS

A multidisciplinary team, including the student's parents or guardian and often the student herself, is responsible for setting goals for students who need special services and have been formally referred. When gathering data on which to base a student's educational program, the team will review the results of various evaluations. These evaluations include information that has been gathered by educational specialists and related service professionals to determine the student's current level of functioning. These data may include reports from:

Formal sources of information for goal setting.

1. *School psychology:* scores, for example, on the Wechsler Intelligence Scale for Children–IV (Wechsler, 2003), Bayley Scales of Infant Development–II (Bayley, 1993), Childhood Autism Rating Scale (Schopler, Reichler, & Renner, 1988), or Kaufman Assessment Battery for Children, 2nd ed. (Kaufman & Kaufman, 2007).
2. *Education:* scores from the Wide Range Achievement Test–Revision 3 (Wilkinson, 1993), Key Math Diagnostic Arithmetic Test, Revised-NU (Connolly, 1998), Woodcock Reading Mastery Tests, Revised (Woodcock, 1998), Brigance Diagnostic Inventory of Basic Skills, Revised (Brigance, 1999).
3. *Adaptive behavior:* Scores from the AAMR (American Association on Mental Retardation) Adaptive Behavior Scale–School Edition, 2nd ed. (Lambert, Nihira, & Leland, 1993), or the Vineland Adaptive Behavior Scales, 2nd ed. (Vineland II) (Sparrow, Cicchetti, & Balla, 2007).
4. *Therapeutic services:* results of physical therapy, occupational therapy, speech–language pathology evaluations.
5. *Physical health:* results from neurological, pediatric, vision, and hearing screenings.

Informal sources of information for goal setting.

In addition to these more formal sources, the goal-setting group should also consider parental desires and concerns. Recommendations from previous teachers are considered as well. The social and academic environmental demands of the present classroom, the home, the projected educational placement, or a projected work site should be examined. Based on this accumulated information, the committee proposes a set of educational

goals for the student. An estimate of progress is then included in the long-term objectives prepared for the student's IEP.

For students who have not been identified as having special needs, formulating educational goals does not involve such an extensive accumulation of information. Assessment may be limited to group achievement tests supplemented by informal teacher-made assessments. Goal setting also is constrained by the adopted curriculum. For example, each class of fourth graders in a given school district is usually expected to learn the same things. Under a standard curriculum, all students at a certain grade level are to be instructed in the natural resources of Peru, the excretory system of the earthworm, multiplication of fractions, and reading comprehension. The teacher's task is to translate these goals into reasonable objectives for each member of a particular class, some of whom may already know these things and some of whom lack the basic skills necessary to learn them. The teacher may write behavioral objectives for the class as a whole, giving consideration to the general characteristics of the group. In addition, if the teacher is to help a particular student who is having problems or to teach a reading group that is progressing slowly, that teacher may write additional behavioral objectives to prescribe a course of instruction that will facilitate learning.

Educational goals for individual students must be developed on the basis of evaluation data but should also consider other factors:

1. the student's past and projected rate of development compared with long-range plans for the future
2. the student's presenting physical and communicative capabilities
3. challenging behaviors that must be brought under control because they interfere with learning
4. skills the student lacks for appropriate functioning in the home, school, and social environments
5. the amount of instructional time available to the student within the school day and within the total school experience
6. prerequisites necessary for acquiring new skills
7. the functional utility of the skills (what additional skills may be built on these?)
8. the availability of specialized materials, equipment, or resource personnel (such as a speech–language pathologist or occupational therapist)

Write goals in observable and quantifiable terms.

Because educational goals are projected over long periods of time, they are written in broad terms. For practical application, however, they need to be written in terms that are observable and quantifiable. As you learned in Chapter 1, applied behavior analysts deal only with observable behaviors.

For students who do not have disabilities or who have mild disabilities, goals are needed only for each curriculum area. For very young students or those with severe disabilities, goals should be written in a number of domains of learning:

1. cognitive
2. communication
3. motor
4. social
5. self-help

6. vocational
7. maladaptive behavior

Long-term educational goals for a student with mild disabilities.

Hypothetical long-term goals for Jason, a student with learning problems in math, and for Tanika, a student with severe disabilities, are as follows.

Jason will

Mathematics: Master basic computation facts at the first-grade level.

Social studies: Demonstrate knowledge of the functions of the three branches of the federal government.

Reading: Identify relevant parts of a story he has read.

Science: Demonstrate knowledge of the structure of the solar system.

Language arts: Increase the creative expression of his oral language.

Physical education: Increase his skills in team sports.

Long-term educational goals for a student with severe disabilities.

Jason's general education teacher will be responsible for setting all goals except the one in mathematics. Jason will probably go to a part-time special education class (resource room) for mathematics. Compare Jason's goals with Tanika's.

Tanika will

Cognitive: Categorize objects according to their function.

Communication: Demonstrate increased receptive understanding of functional labels.

Motor: Develop gross-motor capability of her upper extremities.

Social: Participate appropriately in group activities.

Vocational: Complete assembly tasks for a period of at least 1 hour.

Maladaptive behavior: Decrease out-of-seat behavior.

Self-help: Demonstrate the ability to dress independently.

Teachers use these broad goals to create statements of instructional intent (behavioral objectives).

Behavioral objectives are not simply restatements of goals; they break goals into teachable components. Complex goals may generate many objectives. A goal that states a student will learn to play cooperatively with other children, for example, may require individual objectives identifying the need to share, to take turns, and to follow the rules of a game.

Components of a Behavioral Objective

In order to communicate all the necessary information and provide a basis for evaluation, a complete behavioral objective should

1. identify the learner;
2. identify the target behavior;
3. identify the conditions of intervention; and
4. identify criteria for acceptable performance.

Identify the Learner

Use the student' name.

Behavioral objectives were initially designed to promote individualization of instruction (Gagne, 1985). To promote individualization, the teacher must reidentify the specific student or students for whom each objective is developed. Restatement reinforces the teacher's focus on the individual learner and communicates this focus to others. Thus, we include in a behavioral objective statements such as

- John will. . .
- The fourth graders will. . .
- The participants in the training program will. . .
- The members of the Rappers cooperative learning team will. . .

Identify the Target Behavior

State what the student will do.

After the team selects and defines deficient or excessive target behaviors, the teacher identifies exactly what the student will be doing when the desired change has been achieved. This statement spells out a precise response that is representative of the target behavior.

There are three basic purposes for including this component in the behavioral objective:

1. It ensures that the teacher is consistently observing the same behavior. The observation and recording of the occurrence or nonoccurrence of exactly the same behavior allows for an accurate and consistent reflection of the behavior in the data to be collected.
2. The statement of the target behavior allows for confirmation by a third party that the change observed by the teacher has actually occurred.
3. The precise definition of the target behavior facilitates continuity of instruction when people other than the teacher are involved.

To achieve these three purposes, the target behavior must be described so that its occurrence is verifiable. Precise description minimizes differing interpretations of the same behavior. A student's performance of a given behavior can best be verified when the teacher can see or hear the behavior or see or hear a direct product of the behavior. To attain this precision and clarity in an objective, the verb used to delineate the behavioral response should describe a behavior that is directly *observable, measurable,* and *repeatable.*

Though teachers of the gifted would like students to "discover" and art teachers would like students to "appreciate," objectives described in this manner are open to numerous interpretations. For example, it would be difficult for a third party to decide whether a student had performed the following behaviors:

- *recognize* the difference between big and little
- *understand* the value of coins
- *develop* an appreciation of Melville
- *remain* on task during group work
- *refrain* from aggression

The use of such vague terms leads to confusion and to disagreement about whether a behavior is occurring. Because any behavior can be described in a number of ways, everyone involved in a behavior-change project must agree upon a common description of the behavior. This description is the **operational definition**

of the behavior. It is the definition under which everyone will operate when discussing, observing, counting, reporting, or consulting about this student's performance of this behavior, thus eliminating as much ambiguity as possible. The operational definition contains an agreed-upon description of observable and measurable characteristics of the motor performance of the behavior. These characteristics are clearly stated so that everyone can agree that the behavior has been or has not been performed.

A variety of approaches can be taken to operationally define a behavior. Table 2–1 presents examples of several ways to define on-task behavior. An operational definition usually contains a list of categories or specific examples of behaviors. Fairbanks, Sugai, Guardino, and Lathrop (2007) used a short list of categories, and Umbreit, Lane, and Dejud (2004) used a short list of representative behaviors. The list of broader behaviors would require additional specificity for data collection. Regan, Mastropieri, and Scruggs (2005) and Allday and Pakurar (2007) used more extensive lists with examples. Within their definition, Regan et al. additionally defined a behavior with a negative example.

TABLE 2–1
Operational definitions of on-task behavior.

Fairbanks, Sugai, Guardino, & Lathrop (2007)	Orientation toward the task at hand, compliance with all directions, and working with appropriate materials
Umbreit, Lane, & Dejud (2004)	Looking at the materials or teacher as requested, writing numbers or words related to the assigned task, and complying with instruction
Regan, Mastropieri, & Scruggs (2005)	Student (a) is in designated area of room; (b) is manually engaged with appropriate materials; (c) is reading/writing the question/entry; (d) refrains from making derogatory comments about task/other; (e) asks relevant question(s) to adult; (f) maintains focus on appropriate task and/or the journaling tools, and (g) may appear in thought by intermittently and quietly looking away from material and not writing (engaged only with self).
Allday & Pakurar (2007)	When he or she was (a) actively listening to teacher instructions, defined as being oriented toward the teacher or task and responding verbally (e.g., asking questions about the instructions) or nonverbally (e.g., nodding); (b) following the teacher's instructions; (c) orienting appropriately toward the teacher or task; or (d) seeking help in the proper manner (e.g., raising hand).
Callahan & Rademacher (1999)	Attending appropriately to the instructional task is scored when Seth was observed looking at the teacher or relevant task materials (e.g., textbook, worksheet, paper) and/or participating as required during instructional time. The definition was summarized on a piece of paper for Seth. The paper included an age-appropriate graphic and the words "ON TASK MEANS: (1) I'm in my seat, (2) I'm working quietly, and (3) I'm looking at the teacher or my materials."
Boyle & Hughes (1994)	Manual, purposeful involvement with the task, with delays of no longer than 3 seconds between task steps.
Brooks, Todd, Tofflemoyer, & Horner (2003)	For individual seatwork: "Keeping eyes on work, keeping pencil in hand, and working on the assignment quietly." During group instruction: "keeping eyes on speaker, keeping hands free of materials, and following group directions."

Callahan and Rademacher (1999) prepared a list for the observer and a list specifically for the student. In some cases an element of time is included. Boyle and Hughes (1994), for example, imply a rate of behavior. For Brooks, Todd, Tofflemoyer, and Horner (2003), on-task behavior was a concern during individual seatwork and during group instruction. Therefore, an operational definition was written for both instructional formats.

Operational definitions with multiple indicators may make it harder to count accurately the number of times a behavior occurs, making it difficult to know when the student has reached the criterion. One way to avoid this potential problem is to operationally define the outcome of a complex behavior. To measure on-task behavior, for example, the objective may indicate the number of math problems to be completed within a time limit. The student can accomplish this outcome only by remaining on task. (This difficulty is discussed further in Chapter 3.)

Aggression is an example of a general description of behavior that may be operationally defined both functionally—in terms of its consequences or outcomes—or topographically—in terms of the movements comprising the behavior (Barlow & Hersen, 1984). The authors define aggression functionally as "an act whose goal response is injury to an organism." Finkel, Derby, Weber, and McLaughlin (2003) used this approach by defining aggressive behavior as "any aberrant behavior that involved making physical contact with others in an attempt to injure" (p. 113). Aggression has been defined topographically as hitting, pinching, or shoving peers (Singh, Lancioni, Joy, et al., 2007); hitting, kicking, or biting the teacher (Lerman, Iwata, Shore, & Kahng, 1996); hitting, biting, kicking, slapping, pushing, and shoving directed at the mother (Singh, Lancioni, Winton, et al., 2006); hitting with an open hand or fist, kicking, pinching, biting, or pulling hair (Johnson, McComas, Thompson, & Symons, 2004); hitting, kicking, pushing, pinching, scratching, biting, pulling hair or clothes, or other forms of inappropriate physical contact with a peer or adult; and throwing objects at a person (Lien-Thorne & Kamps, 2005). Providing specific examples of the target behavior increases clarity.

The need for an operational definition is reduced when more precise verbs are used in the objective. Increased precision also promotes more accurate recording of data. A precise behavioral description such as "will sort" rather than "will discriminate," "will circle" rather than "will identify," or "will state orally" rather than "will know" is less likely to be interpreted differently by different observers and reduces the need for repeated verbal or written clarification. Here are some more examples of precise behavioral descriptions:

- will point to the largest item in an array
- will verbally count the equivalent in dimes
- will write a translation of the prologue to *The Canterbury Tales*
- will look at his book or the speaker

One guide for selecting appropriate verbs has been offered by Deno and Jenkins (1967). Their classification of verbs is based on agreement of occurrence between independent classroom observers. They arrived at the three sets of verbs shown in Table 2–2 categorized as *directly observable action verbs, ambiguous action verbs, and not directly observable action verbs*.

In order to evaluate a description of a target behavior, Morris (1976, p. 19) suggests using his IBSO (Is the Behavior Specific and Objective?) test questions:

1. Can you count the number of times the behavior occurs in, for example, a 15-minute period, a 1-hour period, or 1 day? Or, can you count the number of minutes it takes

TABLE 2–2
Observability classification of verbs.

Action Verbs That Are Directly Observable		
to cover with a card	to draw	to place
to mark	to lever press	to cross out
to underline	to point to	to circle
to repeat orally	to walk	to say
to write	to count orally	to read orally
to shade	to put on	to name
to fill in	to number	to state
to remove	to label	to tell what
Ambiguous Action Verbs		
to identify in writing	to check	to construct
to match	to take away	to make
to arrange	to finish	to read
to play	to locate	to connect
to give	to reject	to select
to choose	to subtract	to change
to use	to divide	to perform
to total	to add	to order
to measure	to regroup	to supply
to demonstrate	to group	to multiply
to round off	to average	to complete
to inquire	to utilize	to summarize
to acknowledge	to find	to borrow
to see	to convert	to identify
Action Verbs That Are Not Directly Observable		
to distinguish	to be curious	to solve
to conclude	to apply	to deduce
to develop	to feel	to test
to concentrate	to determine	to perceive
to generate	to think	to create
to think critically	to discriminate	to learn
to recognize	to appreciate	to discover
to be aware	to become competent	to know
to infer	to wonder	to like
to realize fully	to analyze	to understand

Note: From *Evaluating Preplanned Curriculum Objectives,* by S. Deno & J. Jenkins, 1967, Philadelphia: Research for Better Schools. Reprinted by permission.

for the child to perform the behavior? That is, can you tell someone the behavior occurred X number of times or for X number of minutes today? (Your answer should be yes.)

2. Will a stranger know exactly what to look for when you tell him/her the target behavior you are planning to modify? That is, can you actually see the child performing the behavior when it occurs? (Your answer should be yes.)

3. Can you break down the target behavior into smaller components, each of which is more specific and observable than the original target behavior? (Your answer should be no.)

Identify the Conditions of Intervention

Conditions are antecedent stimuli related to the target behavior.

The third component of a behavioral objective is the statement of conditions. The statement of conditions lists antecedent stimuli, including instructions, materials, and setting. It may also include the types of assistance available to the students. These elements may be part of the natural environment in which the behavior is to be performed or they may be provided by the teacher as part of a specific learning task. The statement of conditions helps assure that all aspects of the learning experience will be consistently reproduced.

The teacher may set the occasion for an appropriate response using any or all of several categories of antecedent stimuli:

1. Verbal requests or instructions:
   Sam, point to the little car.
   Debbie, add these numbers.
   Jody, go back to your desk.
2. Written instructions or format:
   Diagram these sentences.
   Find the products.
   Draw a line from each word to its definition.
3. Demonstration:
   This is how you use litmus paper.
   This is how to operate. . .
4. Materials to be used:
   A worksheet with 20 single-digit addition problems
   A tape recorder with the "play" button colored green and the "stop" button colored red
5. Environmental setting or timing:
   In the vocational workshop
   In the cafeteria
   On the playground
   During independent study period
   During transition between classes
6. Manner of assistance:
   Independently
   With the aid of a number line
   With partial physical assistance from the teacher
   Receiving only verbal prompts

Providing appropriate antecedent stimuli will be discussed in Chapter 9.

The teacher must be sure that the verbal or visual cue planned does in fact provide an opportunity for the desired response by the student. That is, a teacher should deliver an unambiguous request or instruction to the student. The teacher who holds a flash card with the word *get* and says, "Give me a sentence for *get*," is likely to hear, "I for get my milk money" or "I for get my homework."

The materials described in the objective should ensure stimulus consistency for the learner and reduce the chance for inadvertent, subtle changes in the learning performance being requested. For example, presenting a red, a blue, and a green sock

and asking the student to "point to red" is a less complex task than presenting a red car, a blue sock, and a green cup and making the same request. Giving the student a page with written instructions to fill in the blanks in sentences is less complex when a list of words that includes the answers is provided. Asking the student to write a story based on a stimulus picture is different from asking the student to write a story without visual stimulus.

The following are examples of condition statement formats:

- Given an array of materials containing . . .
- Given a textbook containing 25 division problems with single-digit divisors . . .
- Given the manual sign for "toilet" . . .
- Given the use of a thesaurus and written instructions . . .
- Given a pullover sweater with the label cued red and the verbal cue "Put on your sweater" . . .
- Given a ditto sheet with 20 problems containing improper fractions having unlike denominators and the written instruction to "Find the products" . . .
- Without the aid of . . .

Careful statement of the conditions under which the behavior is to be performed may prevent problems like the one encountered by Ms. Samuels in the vignette that follows.

Ms. Samuels Teaches Long Division

Ms. Samuels was once again in trouble with a general education teacher. She and Mr. Watson, the sixth-grade math teacher, had agreed she would work on long division with Harvey. Ms. Samuels had carefully checked to be sure the method she taught Harvey to use in the resource room was the same as Mr. Watson's. She made dozens of practice worksheets, and Harvey worked long-division problems until he could do them in his sleep.

Ms. Samuels was predictably horrified then when Mr. Watson asked her if she ever planned to start working on division with Harvey. Investigation revealed that in the general education classroom, Harvey was expected to copy the problems from the math book to notebook paper; he made so many copying errors, he seldom got the correct answer. The conditions under which the task was to be performed were thus significantly different.

As part of a plan of instruction for students with learning problems, teachers may need to include extra support in the form of supplementary cues, such as a model of a completed long-division problem for the student to keep at his desk. It is important to include a description of such supplementary cues in the condition component of the behavioral objective to avoid misunderstandings. When the cue is no longer needed, the objective may be rewritten.

Identify Criteria for Acceptable Performance

Criterion statements set minimum performance standards.

In the criterion statement included in a behavioral objective, the teacher sets the standard for minimally acceptable performance. This statement indicates the level of performance the student will be able to achieve as a result of the intervention. The performance itself has been defined; the criterion sets the standard for evaluation. Throughout the intervention process, this criterion is used to measure the effectiveness of the intervention strategy selected to meet the behavioral objective.

The basic criterion statement for initial learning or **acquisition** indicates the *accuracy* of a response or the response's *frequency of occurrence.* Such statements are written in terms of the number of correct responses, the student's accuracy on trial presentations,

the percentage of accurate responses, or some performance within an error of limitation. Here are some sample criterion statements:

17 out of 20 correct responses

label all 10 objects correctly

with 80% accuracy

on 80% of opportunities

20 problems must be answered correctly (100% accuracy)

4 out of 5 trials correct

on 5 consecutive trials

complete all steps in the toilet-training program independently

list all four of the main characters in a book report of no less than 250 words with no more than 5 errors in spelling

on each occasion

See Chapter 3 for an extended discussion of duration and latency.

Two additional types of criteria may be included when time is a critical dimension of the behavior. *Duration* is a statement of the length of time the student performs the behavior. *Latency* is a statement of the length of time that elapses before the student begins performing the behavior.

- Criterion statements addressing duration:

  will complete within 1 hour

  for at least 20 minutes

  for no more than 1/2 hour

  will return within 10 minutes

  within 2 weeks

- Statements addressing latency:

  within 10 seconds after the flash card is presented

  within 1 minute after a verbal request

Certain types of content require particular criterion levels. When a student is acquiring basic skills on which other skills will be built, a criterion of 80% may not be high enough. For example, learning "almost all" of the multiplication facts may result in a student's going through life never knowing what 8×7 is. There are other skills as well that require 100% accuracy. Remembering to look both ways before crossing the street only 90% of the time may result in premature termination of the opportunity for future learning!

Writing criterion statements in terms of percentage requires care. How many problems, for example, would a student have to solve correctly to satisfy a 90% criterion on a 5-item math quiz?

For certain students, a disability may influence the force, direction, or duration of the criterion set by the teacher. For example, a student may not be able to hammer a nail all the way into a piece of wood; range-of-motion limitations may influence motor capability for reaching; hypotonic muscles (those with less-than-normal tone) may limit the duration of walking or sitting; or a muscular condition may limit the perfection of cursive handwriting.

When setting criteria for acceptable performance, teachers must be careful to set goals that are sufficiently ambitious, yet reasonable. Selection should be based on the nature of the content, the abilities of the students, and the learning opportunities to be provided. Criteria should provide for the development of a functional skill. There is no sense in teaching a student to play a game only so well that he gets beaten every time

he plays or in teaching a student to do math problems only well enough to earn a high F in the general education classroom. There is evidence (Fuchs, Fuchs, & Deno, 1988) that setting ambitious goals results in more learning, but teachers should not set unobtainable goals that will result in frustration for students.

In addition to considering the number or percentage correct and the accuracy of response, writers of behavioral objectives must also determine the number of times a student must meet a criterion to demonstrate mastery. For example, how often must Jane perform a behavior successfully on 8 out of 10 trials before the teacher will be convinced of mastery and allow her to move on to the next level of learning or to the next behavioral objective?

It may be inferred from an open-ended criterion statement that the first time a student reaches 85% accuracy, the skill will be considered "learned" or that from now until the end of the school year, the teacher will continually test and retest to substantiate the 85% accuracy. Either inference could be false. Therefore, a statement such as one of the following should be included in the behavioral objective to provide a point of closure and terminal review:

85% accuracy *for 4 consecutive sessions*

85% accuracy *for 3 out of 4 days*

on 8 out of 10 trials *for 3 consecutive teaching sessions*

will return within 10 minutes on *3 consecutive trips to the bathroom*

Format for a Behavioral Objective

A management aid for the teacher in writing behavioral objectives is the adoption of a standard format. A consistent format helps the teacher include all the components necessary for communicating all intended information. No single format is necessarily superior to others; teachers should simply find one that is compatible with his or her writing style or with administrative policy. Here are two such formats.

Format 1

Conditions: Given 20 flash cards with preprimer sight words and the instruction "Read these words."

Student: Sam

Behavior: will read the words orally

Criterion: within 2 seconds for each word with 90% accuracy on 3 consecutive trials.

Format 2

Student: Marvin

Behavior: will write 20 fourth-grade spelling words in cursive handwriting

Conditions: from dictation by the resource teacher with no more than 2 errors for 3 consecutive weeks.

The following behavioral objectives may be derived from the educational goals previously set for students Jason and Tanika.

Mathematics

Goal: Jason will master basic computation facts at the first-grade level.

Objective: Given a worksheet of 20 single-digit addition problems in the form 6 + 2 and the written instruction "Find the sums," Jason will complete all problems with 90% accuracy for 3 consecutive math sessions.

Social Studies

Goal: Jason will demonstrate knowledge of the functions of the 3 branches of the federal government.

Objective: After reading pages 23–26 in the text *Our American Heritage*, Jason will list the 10-step sequential process by which a bill becomes a law. This list will have no more than one error of sequence and one error of omission. This will be successfully accomplished on an in-class exercise and on the unit-end test.

Reading

Goal: Jason will be able to identify relevant parts of a story he has read.

Objective: Given the short story "The Necklace," Jason will write a minimum 200-word paper that (1) lists all the main characters and (2) lists the sequence of main events, with no more than 2 errors.

Science

Goal: Jason will demonstrate knowledge of the structure of the solar system.

Objective: Given a map of the solar system, Jason will label each planet in its proper position from the sun with 100% accuracy on 2 consecutive sessions.

Language Arts

Goal: Jason will increase the creative expression of his oral language.

Objective: Given an array of photos of people, objects, and locations, Jason will tell a 5-minute story to the class that makes use of a minimum of 7 items, on 3 out of 5 days.

Physical Education

Goal: Jason will increase his skills in team sports.

Objective: Given a basketball, Jason will throw the ball into the hoop from a distance of 10 feet, 8 out of 10 trials for 4 consecutive gym classes.

Recall from our earlier discussion that although Jason has mild learning problems, Tanika has much more severe disabilities. Here are some objectives and corresponding goals for Tanika.

Cognitive

Goal: Tanika will be able to categorize objects according to their function.

Objective: Given 12 Peabody cards (4 foods, 4 clothing, 4 grooming aids), a sample stimulus card of each category, and the verbal cue "Where does this one go?", Tanika will place the cards on the appropriate category pile with 100% accuracy, for 17 out of 20 trials.

Communication

Goal: Tanika will demonstrate increased receptive understanding of functional labels.

Objective: Given an array of 3 objects found in her snack-time environment (cup, spoon, fork) and the verbal cue "Pick up the . . .," Tanika will hand the teacher the named object 9 out of 10 times for four consecutive snack times.

Motor

Goal: Tanika will develop gross motor capability of her upper extremities.

Objective: Given a soft rubber ball suspended from the ceiling and the verbal cue "Hit the ball," Tanika will hit the ball causing movement 10 out of 10 times for 5 consecutive days.

Social

Goal: Tanika will learn to participate appropriately in group activities.

Objective: When sitting with the teacher and two other students during story time, Tanika will make an appropriate motor or verbal response to each of the teacher's questions when called upon a minimum of 3 times in a 10-minute period for 5 consecutive days.

Self-Help

Goal: Tanika will demonstrate the ability to dress herself independently.

Objective: Given a pullover sweater with the back label color-cued red and the verbal cue "Put on your sweater," Tanika will successfully complete all steps of the task without physical assistance 2 out of 3 trials for 4 consecutive days.

Vocational

Goal: Tanika will complete assembly tasks for a period of at least 1 hour.

Objective: Given the four parts of a plumbing "U" in sequential order, Tanika will assemble at a rate of one per 3 minutes without error during 3 vocational periods for 4 weeks.

Maladaptive Behavior

Goal: Tanika will decrease out-of-seat behavior.

Objective: In the period from 9:00 to 9:20 a.m. (functional academics), Tanika will remain in her seat, unless given permission by the teacher to leave, for 5 consecutive days.

Professor Grundy's Class Writes Behavioral Objectives

It was the time of the semester for Professor Grundy's 8 o'clock class to learn about behavioral objectives. After presenting a carefully planned lecture (remarkably similar to the first part of this chapter), Grundy asked if there were any questions. Dawn Tompkins stopped filing her nails long enough to ask, with a deep sigh, "Yes, Professor, would you please tell me what a behavioral objective is, exactly?"

"I was under the impression, young lady, that I had done just that," replied Grundy. "Is anyone else confused?"

A chorus of muttering and rumbling ensued from which Grundy was able to extract clearly only two questions: "Is this covered in the book?" and "Will it be on the test?"

After once more presenting a drastically abbreviated description of the components of a behavioral objective, Grundy announced that each member of the class was to write a behavioral objective for the curriculum area of science and present it to him for checking before leaving class. This announcement, followed by a chorus of groans and considerable paper shuffling, also brought forth a flurry of hands:

"You mean list the components?"
"No," said Grundy. "Write an objective."
"You mean define a behavioral objective?"
"No," said Grundy. "Write one."
"But you never said anything about writing them."
"What," Grundy retorted, "did you think was the purpose of the lecture?"

After everyone who lacked these tools had been provided with paper and pencil, silence descended upon the class. DeWayne was the first one finished and proudly presented his objective to the professor:

To understand the importance of the digestive system.

"Well, DeWayne," said the professor, "that's a start, but do you not remember that a behavioral objective must talk about behavior? Remember the list of verbs I gave you . . ." When DeWayne continued to look blank, Grundy rifled through his briefcase and found a copy (see Table 2–2).

"Look here," the professor said, "use one of these directly observable verbs."

DeWayne returned some time later with his rewritten objective:

To label the parts of the digestive system.

"Good, DeWayne," sighed the professor, "that's a behavior, all right. Now, do you recall the components of a behavioral objective?" Once again, DeWayne looked blank. Grundy carefully wrote:

Conditions Student Behavior Criteria

across the long side of the automatic teller receipt DeWayne had evidently found in his wallet. (Grundy was not surprised to note that the receipt indicated a negative balance.) DeWayne returned to his desk.

An hour and a half later, as Grundy was regretting ever having made this assignment, DeWayne returned again:

Given an unlabeled diagram of the human digestive system, fourth-grade students will label the major parts of the digestive system (mouth, esophagus, stomach, small intestine, large intestine) with no errors.

Grundy read DeWayne's objective with interest, because his own digestive system was beginning to be the major focus of his attention. "Excellent, DeWayne," said the professor. "I suppose it's too late to get lunch in the cafeteria. Why didn't you do this in the first place?"

"Well, Professor," answered DeWayne, "I didn't really understand what you wanted. I'm still not sure I could do another one."

After getting some crackers from a vending machine, Grundy returned meditatively to his office. He found a piece of paper and began to write as he munched:

Given a worksheet listing appropriate verbs and the components of a behavioral objective, students enrolled in Education 411 will write five behavioral objectives including all components.

After musing for a few minutes, he added:

in less than half an hour.

"Perhaps," Grundy muttered to himself, "if I had been sure what I wanted and told the students at the front end, they would have had less trouble figuring it out."

Expanding the Scope of the Basic Behavioral Objective

Once a student or group of students has acquired the behavior described in an objective, teachers may simply note that the objective has been mastered and move on to the next one in the sequence. This may be inappropriate unless the student can perform the behavior in circumstances different from the initial teaching environment. In order for the students to have functional behaviors, those that can be performed under different conditions, to different criteria, or in the absence of reinforcement contingencies, provision must be made to expand students' ability to use the behavior. Two possible perspectives on expanded use are

1. programming according to a hierarchy of response competence.
2. programming according to a hierarchy of levels of learning.

Hierarchy of Response Competence

A measure of response accuracy (8 out of 10 correct, for example) is only one dimension for evaluating performance. It represents the acquisition level of response competence. At this level, we merely verify the presence of the ability to do something the student was not previously able to do and the ability to do it with some degree of accuracy. Moving to measures of competence in performance beyond accuracy, beyond this acquisition level, requires alterations or additions to the statements of criteria and conditions. Such alterations reflect a hierarchy of response competence. Once a child can perform the behavior, we are then concerned with fluency, or rate, of performance, as well as performance under conditions other than those imposed during the initial teaching process.

A response hierarchy should contain the minimum levels of acquisition, fluency, maintenance, and generalization.

As an example of the use of this hierarchy, let us assume John has reached the acquisition level on the following objective:

> ***Given two quarters, two dimes, two nickels, and one penny and the verbal cue "John, give me your bus fare," he will hand the teacher coins equaling 75 cents 8 out of 10 trials for three consecutive sessions.***

Lauren has reached acquisition on this objective:

> ***Given a worksheet with 20 division problems with two-digit dividends and single-digit divisors, Lauren will write the correct answer in the appropriate place on the radical with 90% accuracy for 4 consecutive days.***

After John and Lauren have met these stated criteria for their performances, the teaching concern should turn to their **fluency** of performance, or the rate at which they perform the behavior. Fluency refers to the appropriateness of the rate at which the student is accurately performing this newly acquired response. In John's case, we know that he can select the appropriate coins to make 75 cents, but this does him little good if when we take him to the bus, it takes him 5 minutes to do it. The bus driver cannot wait this long. In Lauren's case, we know she can now solve division problems, but it takes her so long that either we interrupt her when her reading group is scheduled or she misses part of her reading lesson so she can finish her problems.

In both instances, the students are demonstrating accurate performance at an inappropriate rate. Recognizing the necessity for an appropriate rate of performance, a teacher can indicate an acceptable fluency when the behavioral objective is written. This is accomplished by adding a time limit to the statement of criteria, as found in parentheses in the following objectives:

> ***Given two quarters, two dimes, two nickels, and one penny and the verbal cue "John, give me your bus fare," he will hand the teacher coins equaling 75 cents (within 30 seconds) 8 out of 10 trials for three consecutive sessions.***

Given a worksheet with 20 division problems with two-digit dividends and single-digit divisors, Lauren will write the correct answer in the appropriate place on the radical (within 20 minutes) with 90% accuracy for 4 consecutive days.

For typical learners and those with mild disabilities, the rate is often included in the initial objective, thus combining acquisition and fluency in a single instructional procedure. Instructional attention is given to fluency because when a student's performance becomes fluent, the behavior is retained longer, persists during long periods on the task, is less affected by distractions, and is more likely to be available in new learning situations (Pierce & Cheney, 2004).

Programming for maintenance will be discussed in Chapter 10.

It is not necessary to adjust the original behavioral objective to include the level of competence labeled **maintenance.** Maintenance is the ability to perform a response over time without reteaching. Maintenance-level competence is confirmed by using postchecks or probes, during which the teacher rechecks the skill to be sure the student can still do it. Maintenance may be promoted through building in the opportunity for overlearning trials and distributed practice. **Overlearning** refers to repeated practice after an objective has been initially accomplished. An optimum number of overlearning opportunities is approximately 50% of the number of trials required for acquisition of the behavior. If it takes John 10 teaching sessions to learn to tie his shoes, we should ideally provide 5 additional sessions for overlearning. *Distributed practice* is practice that is spread out over time (as opposed to *massed practice,* which is compressed in time). An example of massed practice familiar to college students is cramming for an exam. The material may be learned between 10 and 6 the night before the test, but most of it will be rapidly forgotten. If maintenance is desired, the preferable approach is studying for short periods every evening for several weeks before the exam, using distributed practice. Another means of providing for maintenance, alteration of schedules of reinforcement (Skinner, 1968), will be discussed in Chapter 7.

Distributed practice is a more efficient way of learning for long-term maintenance.

The level of response competence labeled **generalization** is of great importance in assuring that a behavior is functional. A student has a generalized response if she can perform—and adapt, if necessary—the behavior under conditions different from those in place during acquisition. A generalized response is one that also continues to occur after instruction has been terminated. A response may be generalized across at least four basic dimensions. The condition statement may be written to reflect the student's ability to perform the behavior in response to various verbal or written instructions, with various materials, for or with various persons, and in various environments (settings). The following examples illustrate this point.

Various Instructions

Given an array of coins and the verbal instruction, "Give me bus fare" ("Give me 75 cents," "Give me what you need for the bus") . . .

Given a worksheet with 30 one-digit subtraction problems and the verbal (or written) instruction, "Find the difference" ("Solve these problems," "Write the answers to these problems") . . .

Various Materials

Will write his name, address, phone number, and birth date in the appropriate blanks on at least three different job application forms . . .

Will demonstrate the multiplicative principle of math using counting chips (a number line, paper and pencil) . . .

Various Persons

Will use the sign for "toilet" as a signal of need to her teacher (parent) . . .

Will comply with instructions from his math (English, social studies, science) teacher (mother, father, coach, piano teacher) . . .

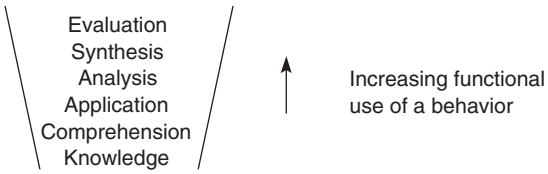
Various Settings

Will pull up his pants after toileting in the restroom in the special education class (in the restroom near the class he joins for art) . . .

Will remain in his seat and complete assignments in math (English, social studies, science) class . . .

Hierarchy of Levels of Learning

It may seem that writing behavioral objectives inevitably focuses teacher attention on concrete, simple forms of learning. Indeed, this has been one of the most frequent criticisms of a behavioral approach. It is not necessary, however, to confine behavioral objectives to lower levels of learning. Bloom (1956) has proposed hierarchies of learning in cognitive, affective, and psychomotor areas. These hierarchies classify possible learning outcomes in terms of increasingly abstract levels. They are helpful in writing objectives in behavioral terms because they suggest observable, measurable behaviors that may occur as the result of both simple and complex learning. The cognitive hierarchy, which will serve as our example, contains six levels of learning, as shown in the diagram (Bloom, 1956).



Many behavioral objectives are written in terms of the knowledge level of the hierarchy—we simply want students to demonstrate they know or remember something we have taught them. Once the student has achieved mastery on the lowest of the six levels, the teacher can shift programming toward higher levels of learning by preparing subsequent objectives that alter the target behavior and criterion statements. As an aid in this process, Gronlund (1985) prepared a table (Table 2–3) that illustrates behavioral terms appropriate to describe target behaviors at each level of learning.

Knowledge

Bloom (1956) defines learning at the knowledge level as the recall or recognition of information ranging from specific facts to complete theories. These memory functions are the only behavior to be demonstrated at this basic level of cognitive learning. The following acquisition objectives are examples written for students at this level:

> ***After reading*** Biology for Your Understanding ***and completing the exercise in Chapter 2, Virginia will list the biological categories of the Linnaean system in their order of evolutional complexity without error during two class sessions and on a unit-end exam.***

TABLE 2–3
Examples of general instructional objectives and behavioral terms for the cognitive domain of the taxonomy.

Illustrative General Instructional Objectives	Illustrative Behavioral Terms for Stating Specific Learning Outcomes
Knows common terms Knows specific facts Knows methods and procedures Knows basic concepts Knows principles	Defines, describes, identifies, labels, lists, matches, names, outlines, reproduces, selects, states
Understands facts and principles	Converts, defends, distinguishes, estimates, explains, extends, generalizes, gives examples, infers, paraphrases, predicts, rewrites, summarizes
Interprets verbal material Interprets charts and graphs Translates verbal material to mathematical formulas Estimates future consequences implied in data Justifies methods and procedures	
Applies concepts and principles to new situations Applies laws and theories to practical situations Solves mathematical problems Constructs charts and graphs Demonstrates correct usage of a method or procedure	Changes, computes, demonstrates, discovers, manipulates, modifies, operates, predicts, prepares, produces, relates, shows, solves, uses
Recognizes unstated assumptions Recognizes logical fallacies in reasoning Distinguishes between facts and inferences	Breaks down, diagrams, differentiates, discriminates, distinguishes, identifies, illustrates, infers, outlines, points out, relates, selects, separates, subdivides
Evaluates the relevancy of data Analyzes the organizational structure of a work (art, music, writing)	
Writes a well-organized theme Gives a well-organized speech Writes a creative short story (or poem, or music) Proposes a plan for an experiment	Categorizes, combines, compiles, composes, creates, devises, designs, explains, generates, modifies, organizes, plans, rearranges, reconstructs, relates, reorganizes, revises, rewrites, summarizes, tells, writes
Integrates learning from different areas into a plan for solving a problem Formulates a new scheme for classifying objects (or events, or ideas)	
Judges the logical consistency of written material Judges the adequacy with which conclusions are supported by data Judges the value of a work (art, music, writing) by use of internal criteria Judges the value of a work (art, music, writing) by use of external standards of excellence	Appraises, compares, concludes, contrasts, criticizes, describes, discriminates, explains, justifies, interprets, relates, summarizes, supports

Note: From *How to Write and Use Instructional Objectives* (5th ed.), by Norman E. Gronlund, 1991, Upper Saddle River, NJ: Prentice Hall, Inc. Copyright © 1991. Reprinted by permission.

Given the symbols for the arithmetic processes of addition, subtraction, multiplication, and division, Danny will respond with 90% accuracy on a multiple-choice test of their labels and basic functions.

Given a list of Shakespearean plays, Deborah will underline the names of the tragedies with no more than one error.

Comprehension

Once the student has reached the performance criterion at the knowledge level, the teacher moves to the comprehension level, the understanding of meaning. The student may demonstrate comprehension by paraphrasing and providing examples.

Here are some sample objectives at this level:

Given the Linnaean system of biological classification, Virginia will provide a written description of an organism in each category. The description will include at least one factor that distinguishes the category from others.

Given a worksheet of 40 basic arithmetic examples requiring addition, subtraction, multiplication, and division, Danny will complete the sheet with 90% accuracy.

Given the metaphoric passage, "Oh that this too, too solid flesh would melt . . ." from Hamlet*, Deborah will write an essay describing the literal intent of the passage. The essay will be a minimum of 300 words.*

Application

Programming at Bloom's *application level* requires the student to use the method, concept, or theory in various concrete situations. Consider these objectives:

Given the names of five organisms and the Linnaean system, Virginia will place each in its proper category and write a list of rationales for placement. Each rationale will contain a minimum of two reasons for placement.

Given a set of 10 paragraphs that present problems requiring an arithmetic computation for solution, Danny will write the correct answer, showing all computations with 100% accuracy.

After reading Hamlet, *Deborah will be able to explain the parallels between Hamlet's ethical dilemma and the problem of abortion and to cite an additional current parallel example of her own choosing.*

Analysis

Analysis is the ability to break down material into its constituent parts in order to identify these parts, discuss their interrelationship, and understand their organization as a whole. The following objectives are analytically oriented:

Given a list of five organisms, Virginia will use appropriate references in the library to investigate and report to the class the role of the organisms in either the food chain or in the ecological stability of their habitat.

Given a written statement of the associative property, Danny will be able to explain accurately to the class, using examples at the chalkboard, the property's relation to the basic additive and multiplicative functions.

After having read Hamlet *or* Macbeth, *Deborah will guide the class in a discussion of the play's plot development. This discussion will be based upon a schematic representation of each scene that she will provide in written form.*

Synthesis

At the cognitive level of *synthesis,* the student should demonstrate the ability to bring parts together, resulting in a different, original, or creative whole:

> ***Given a list of reference texts, Virginia will write a 1,000-word summary explaining the biological classifications in Darwin's theory of evolution. The paper will be evaluated on the basis of accuracy, completeness, organization, and clarity.***
>
> ***Given the numerical systems of base 10 and base 2, Danny will orally demonstrate the use of the functions of addition, subtraction, multiplication, and division within each system.***
>
> ***Given the study of the Shakespearean tragedy*** Macbeth, ***Deborah will rewrite the end of the play in iambic pentameter, assuming that the murder of the king was unsuccessful.***

Evaluation

The highest level of learning demonstrated in this hierarchy is *evaluation.* The student is asked to make a judgment of value:

> ***Based on the principles of mutual exclusion, Virginia will devise a taxonomy for the classification of means of transportation and provide a justification for the categories created and their constituent parts.***
>
> ***Given a set of unknown values and a given arithmetic computational function, Danny will explain the probability of differing answers that may be correct.***
>
> ***Given plays by Shakespeare and Bacon, Deborah will state a preference for one and justify her preference in a 500-word essay based on some element(s) of style.***

Learning Levels for the Learner with Limitations

In most instances of planning for expanded instructional intent, we tend to focus on a hierarchy-of-response competence for learners with significant disabilities and a hierarchy of levels of learning for the typical or above-average learner. This dichotomy is not necessarily warranted simply by the level of the student's functioning. Consider the following examples of how we may write behavioral objectives for the limited learner in conjunction with levels of learning:

Knowledge:	Given a common coin and the verbal cue, "What is the name of this?," George will state the appropriate label on 18 out of 20 trials for five consecutive sessions.
Comprehension:	Given a common coin and the verbal cue "What is this worth?", George will count out the coin's equivalent in pennies and state something to the effect that "A dime is worth 10 pennies" on 8 out of 10 trials for each coin.
Application:	When presented with 10 pictures of food items, each with its cost written on it, George will count out coins equal to the amount written upon the verbal cue "Show me the amount" on 18 out of 20 trials.
Analysis:	When presented with pictures of items, each with its cost printed on it, a $1 bill, and a verbal cue such as "Can you buy a pencil and a newspaper?" George will respond correctly on 18 out of 20 trials.

Synthesis:	Given a $1 bill and the instruction to buy various priced items, George will simulate the buying exchange and decide whether he was given correct change without error on 10 trials.
Evaluation:	Given a $1 bill and a 5-mile ride from the workplace to his home, George will use the $1 for the bus ride rather than a candy bar.

Even learners with limitations can acquire higher-level cognitive skills.

BEHAVIORAL OBJECTIVES AND THE IEP

IDEA requires an IEP for every student with a disability.

The development of educational goals (long-term objectives) and behavioral objectives (short-term objectives) for students in need of special education services was included as one of the mandates of the original Education for All Handicapped Children Act of 1975 (P.L. 94-142), and its current successor, the Individuals with Disabilities Education Improvement Act of 1999 (P.L. 108-446, IDEA). Among the results of this legislation have been the formalization of the planning aspects inherent in the writing of goals and objectives and the provision for active parental participation in the educational planning process. This planning process results in the development of an **individualized education program (IEP).** The IEP has at its core the listing of the goals and objectives for the student's educational program for the year and how progress toward these goals will be measured (Siegel, 2007). In addition to this core element, an IEP contains components or statements regarding transition planning and services, positive behavioral interventions and supports, participation in state and district assessments, extended school year services, participation in the general education curriculum (including necessary modifications), and interaction with students not identified as disabled. The federal rules and regulations include six elements as part of the IEP:

1. a statement of the student's present levels of educational performance, which may be assessed by standardized norm-referenced tests, classroom-based assessments, direct observation, and curriculum-based measures (Yell & Stecker, 2003)
2. a statement of measurable annual goals for students with mild disabilities or a statement of annual goals and short-term instructional objectives for students with significant disabilities
3. appropriate objective criteria and evaluation procedures and schedules for determining, on at least an annual basis, whether the short-term instructional objectives are being achieved
4. a statement of the specific special education and related services to be provided to the student
5. projected dates for initiation of services and the anticipated duration of the services
6. the extent to which the student will be able to participate in general education programs, and any modifications or accommodations necessary to enable that participation

These elements demonstrate a parallel in procedural format between the development of behavioral objectives and the development of the IEP. Both processes include the accumulation of data to determine the student's current levels of performance, the statement of appropriate goals, the development of behavioral objectives (short-term) for attaining the goals, and a review of objective mastery. Neither a behavioral objective nor an IEP short-term objective contains a statement of the instructional methodology to be used for achieving the objective (Bateman & Linden, 1998; Mager, 1997: Yell & Stecker, 2003).

The following recommendations will facilitate management of an IEP and monitoring of its constituent objectives:

1. Short-term objectives should be sequentially related to goal statements. Across a sequence of objectives the teacher can systematically alter the elements of objectives. For example, aspects of the *conditions* (materials used, setting, format, type or amount of assistance or accommodation provided to the student); the *response* (e.g., response mode; the cognitive or physical difficulty or complexity of the response required); and/or the *criterion* (e.g., the amount of response required in number or duration; rate; number or types of errors permitted) may be increased or altered as the teacher shapes the student's response to one that meets a standard set for performance or is more functional. The examples in Figure 2–1 by Lignugaris, Kraft, Marchard-Martella, and Martella (2001) depict a sequence of objectives for a mathematics goal in which the description of the materials in the condition statement increases in difficulty, as well as a sequence for spelling in which there is an increasing pool of words.
2. In the case of students with mild disabilities, the goals and short-term objectives should deal directly with the reason for their referral for special education services. "They need be written only for the special services necessary to meet the child's needs arising from the disability, not for the child's total program, unless all areas are so affected" (Bateman & Linden, 1998, p. 43).
3. For students with moderate, severe, and profound disabilities, two or three short-term objectives per curriculum domain should be included in the IEP because in most cases all areas of the student's educational performance are affected by the disability.
4. New short-term objectives should not be added until maintenance has been achieved on current objectives and generalization instruction has begun.
5. Management of the IEP should be a continuous process. Teachers and administrators should not overlook the regulations stating that a review should be conducted "on at least an annual basis," not "only on an annual basis."
   a. The objectives for students with mild disabilities should be reviewed as soon as achievement has been verified to assess whether the original need for special education services still exists.
   b. Reasonable review dates should be set for objectives of students with moderate and severe disabilities. As objectives are met, the teacher should add new short-term objectives and notify the committee members, including the parents, in writing, with full justification provided at the annual review. Such a procedure will foster the student's progress and prevent stagnation of instruction until the full committee can be gathered.

When Jason's math skills are at grade level, he no longer has a disability.

Tanika's objectives should be reviewed frequently so that she will make the maximum possible progress.

FIGURE 2–1 Sample IEP.

Student: Leon

Short-Term Instructional Objectives	Person Responsible	Criteria for Mastery	Date Reviewed	Mastery Yes No
1. Complete 20 single-digit addition problems in the form 6 + 2, given written instructions to "find the sum"		90% accuracy for 3 consecutive math sessions		

6. Review dates should be set considering the need for instruction at higher levels of learning to promote full functional use of a skill.

In addition to the basic core, two additional planning components are contained in the IEP. These are the individual transition plan and the behavioral intervention plan.

The Individual Transition Plan

The **individual transition plan (ITP)** is that component of the IEP that links long-term planning for a student's postsecondary life with the annual school year planning of the IEP. When a student reaches the age of 16, or earlier if determined appropriate, the IEP team must begin to determine what instructional and educational experiences will assist the student to prepare for transition from secondary education to postsecondary life. The plan must include appropriate measurable postsecondary goals, based on appropriate transition assessments focused on training, employment, education, and independent living. It must also list the specific transition services that will be required to help the student reach the transition goals (Siegel, 2007).

The purpose of the ITP is to help the student, family members, educators, and adult service providers focus on the student's future after completing secondary school. Tentative decisions are made about the student's living arrangements, community involvement, employment, and postsecondary education. These decisions are examined annually to determine the need for revision in light of the student's skill mastery and evolving preferences. The IEP team must determine what, if any, instruction, related services, community experiences, employment, and other adult living objectives are necessary to meet the postsecondary goals. After transition outcomes are identified, more specific IEP goals and objectives are developed annually to ensure that students will be ready for the transition.

The relationship between transition outcome recommendations of the ITP and IEP goals and objectives can be illustrated as follows:

ITP Outcome Recommendation: Obtain a supported employment position in the food-services industry.

IEP Annual Goal: Lana will participate in three community-based vocational instruction sites during the 2007–2008 school year.

IEP Objectives:

1. Given a set of models, Lana will separate glasses, dishes, and silverware with at least 90% accuracy in three work sites on six occasions.
2. Lana will train in three food-service work sites on a task analysis of dishwashing, independently completing 100% of the task steps, at the indirect level of supervision, for a range of 6 to 8 hours a week, for no more than 215 hours.

Transition Outcome Recommendation: Live semi-independently in an apartment with a roommate with supervision provided by Peachtree Residential Supports, Inc.

IEP Annual Goal: Sally will learn to prepare simple meals in a variety of settings.

IEP Objectives:

1. Sally will independently operate a microwave oven to prepare a prepackaged meal at home, at school, and in the breakroom on the vocational training site, 2 days per week for 12 weeks.
2. Given a list (picture/word) of five breakfast items, Sally will locate and purchase at least four of the items in the local grocery once per week during September, November, and January.

Transition Outcome Recommendation: Enrollment in the Fannin County Technical College program in Nursery and Landscape Management.

IEP Annual Goal: Jacob will enroll in the Earth Science course.

IEP Objective: Given the unit on rock formations, Jacob will identify at least 5 of the 8 rock formations, on 4 of 5 opportunities by the end of the grading period.

The Behavioral Intervention Plan

The **behavioral intervention plan (BIP)** is the component of the IEP that attempts to link the management of behaviors that impede instruction with positive intervention strategies to reduce the occurrence of inappropriate behaviors and increase socially and contextually appropriate behavior. In order to make this connection, a BIP includes the following elements:

1. the operational definition of the target behavior impeding instruction.
2. the results of a functional behavior assessment. A functional behavior assessment is conducted to determine the purpose (function) the inappropriate behavior serves for the student. It requires analysis of data collected on the events that precede (antecedents) and follow (consequences) the target behavior (see Chapter 6).
3. the intervention strategies to be used. These are selected based on the function of the inappropriate behavior. When selecting an intervention strategy, the IEP team is required to consider, if appropriate, strategies including positive behavioral interventions and supports to address behaviors that interfere with learning. Such strategies and supports are reinforcement based and may also include curriculum and instructional modifications and changes in the classroom environment.
4. the behavioral objectives that state the agreed-upon change in the form or rate of the inappropriate behavior, or alternative behaviors that serve the function of the inappropriate behavior for the student.

The following is an example of these elements of a BIP (an example of a full BIP appears in Chapter 6).

Operational definition of the impeding behavior: Leon has tantrums during individual math work times. His tantrums consist of banging his hands and arms on the desk, tearing his papers, and yelling "no, no, no."

Perceived function of the inappropriate behavior resulting from a functional behavior assessment: An analysis of the data indicates that Leon engages in tantrums in order to escape from the academic demand placed on him during math work group.

Positive behavior intervention: (a) Differential reinforcement schedule of hand raising. Begin with redirection of the flailing hands and the teacher's asking: "Leon, do you need help with your work or do you need a break? (break defined as 2 minutes with assignment removed from desk); (b) initially reduce the number of math problems he must complete during the 20-minute period; and (c) provide three problems to do first with an assigned classmate at beginning of math work group.

Behavioral objective(s): Given a 20-minute period of individual math work (division problems with two-digit divisors), Leon will raise his hand when (a) needing assistance or (b) wanting a break from work, 100% of instances for 4 weeks.

According to IDEA there are two circumstances that require the preparation of a BIP (Turnbull, Wilcox, Stowe, & Turnbull, 2001):

1. If at the time of the development, review, or revision of an IEP the team is aware of a pattern of behavior that impedes the learning of the student or of another student, a BIP should be developed as part of the IEP.
2. At the time that an occurrence of behavior subjects a student to disciplinary action (that results in a suspension of up to 10 days, or that results in a change of placement, e.g., suspension, expulsion, or removal to an Interim Alternative Educational Setting): (a) if a BIP does not exist, the team must develop one, and (b) if the IEP contains a BIP, it must be reviewed for necessary modifications

SUMMARY

We described the process of writing behavioral objectives and the relationship between such objectives and the IEP required for students with disabilities. This process is an integral part of any program for behavior change, whether the program is directed toward academic or social behavior. A program for changing behavior is unlikely to be successful unless we are sure what constitutes success. Behavioral objectives facilitate communication, so that everyone knows the goal of instruction. They also provide for evaluation, so that everyone knows whether the goal has been reached.

KEY TERMS

behavioral objective
pinpointing
educational goals
operational definition
acquisition
fluency
maintenance
overlearning
generalization
individualized education program (IEP)
individual transition plan (ITP)
behavioral intervention plan (BIP)

DISCUSSION QUESTIONS

1. Most teachers are required to write objectives as a part of lesson planning or in IEPs. Many teachers consider writing objectives as unnecessary paperwork. Does the time taken to write objectives improve instruction, or are those teachers' right?
2. What is the significance of a criterion statement? For what kinds of objectives are criterion statements more or less important ?
3. Replace each of the following vague verbs with one that is more specific:

   Mario will be able to discriminate between a few and a lot.
   Mario will be able to recall the major rivers of the United States.
   Mario will be able to identify the parts of a flower.
   Mario will be able to understand the results of global warming.
   Mario will know the 6 and 8 multiplication tables.
   Nikki will recognize the main characters of a story.
   Nikki will appreciate the differences among various cultures.
   Nikki will be competent in telling time.
   Nikki will appreciate the works of Monet.
   Nikki will learn to operate a calculator.

Chapter 3

Procedures for Collecting Data

Did you know that . . .

Teachers react strongly to suggestions that they collect data in their classrooms:

- "I don't have time to write down everything anyone does I just don't think I can manage shuffling all those sheets of paper, handling stopwatches and wrist counters, and giving proper cues. When am I supposed to concentrate on teaching?"
- "This data collection adds an extra hour a day, at least, in summarizing the data, putting the data on graphs, and so on. Where's that time supposed to come from?"
- "Give me a break."

Chapter Outline

Most teachers regard the kind of data collection procedures that we shall discuss in this chapter with the same enthusiasm they reserve for statistics. In some cases, their comments are thoroughly justified. Some of the systems we will review are not practical for everyday classroom use. Classroom teachers may never use some of the more complex systems. Understanding how these systems work, however, helps in understanding published research about applied behavior analysis. This chapter describes the most common data collection systems and shows how many of them can be adapted for classroom use.

A Rationale

Even after accepting the feasibility of data collection in the classroom, many teachers see little value in it. Beyond recording grades on tests, most teachers have traditionally kept very few records of their students' academic and social behaviors. There are, nevertheless, excellent reasons for teachers to collect classroom data.

First, observation and measurement make it possible to determine very accurately the effects of a particular instructional strategy or intervention. Precise observation and measurement of behavior enable teachers to determine the success or failure of their strategies. Second, the types of data collection procedures discussed in this chapter allow for ongoing (formative) as well as terminal (summative) evaluation of instruction or intervention. The data collected enable teachers to make decisions and alterations during the course of a program rather than waiting, perhaps for weeks or months, to see if it was ultimately successful. Such use of systematic formative evaluation significantly increases students' achievement, both statistically and practically (Fuchs & Fuchs, 1986). Finally, collecting and reporting effect-based data is the ultimate tool of accountability.

Behavioral evaluation requires observation of students' current functioning and ongoing progress.

By writing behavioral objectives, teachers communicate their intent to change particular behaviors. They also state the criteria they will use to judge whether change procedures have been successful. In many classroom situations, the intervention's effect on the students' original level of performance would be evaluated by administering a pretest and posttest. However, the precision desired within a behavioral approach to instruction and in program evaluation necessitates additional data.

Behavioral evaluation has two requirements. The first is a detailed observation of a student's current functioning. This observation should reflect the conditions and description of the behavior stated in the objective. For example, a behavioral objective stating that students should solve 25 long-division problems in 30 minutes requires that the teacher determine how many long-division problems the students can already solve in 30 minutes. Second, evaluation of an instructional program must facilitate ongoing monitoring of the teaching and learning process and provide a system for terminal evaluation. Evaluation must be continuous so that programs can be adjusted as instruction progresses. As the students in our example receive instruction in long division, the teacher might record daily how many problems they solve in 30 minutes, thus providing continuous evaluation. The

monitoring process can provide guidelines for continuing or changing instructional techniques and help avoid false assumptions about student progress. Such false assumptions are unfortunately very common, as illustrated by the following vignette.

Ms. Waller Goes Electronic

Ms. Waller was ecstatic. After months of complaining that she had no materials to use to teach reading to her most challenging reading group, she had received several computers and a program to teach reading. The salesman proudly demonstrated the machine and pointed out the features that justified the hundreds of dollars invested.

"All you have to do," he assured her, "is hook the little, er, students up to these here headphones, drop in a CD-ROM, and turn this baby on. Everything else is taken care of . . . you don't do a thing."

Ms. Waller briskly administered the pretest included in the materials, scheduled each student for 15 minutes a day on the computer, and assumed that her worries were over.

At the end of the school year, Ms. Waller administered the posttest. Imagine her distress when, although several members of the group had made remarkable progress, some students had made none at all.

"I don't understand," she wailed. "The computer was supposed to do everything. How was I supposed to know it wasn't working?"

"Perhaps," suggested her principal, kindly, as he wished her success in her new career as an encyclopedia salesperson, "you should have checked before now."

Choosing a System

Dimensions for observation of behavior.

The first step in the evaluation of ongoing measurement of behavior is the selection of a system of data collection. The characteristics of the system selected must be appropriate to the behavior being observed and to the kind of behavior change desired.

Behavior may be measured and changed on a number of dimensions (White & Haring, 1980).

1. *Frequency:* The frequency of behavior is simply the number of times a student engages in it.
   Brett got out of his seat 6 times in 30 minutes.
   Yao did 6 of 10 *math problems during a timed trial.*
   Marvin had 8 *tantrums Wednesday.*
   Lois' hand was in her mouth 5 times *during storytelling.*

   When determining frequency of occurrence of a behavior, we count the number of times the behavior occurs within an observation period (for example, 10 seconds or a 40-minute science class). If we want to make comparisons of the frequency of a behavior across observation periods (from one lunch period to another), the observation periods should be of the same length.

   If a behavior can occur only a limited number of times, that information should be provided as part of the frequency data. Knowing that Yao solved 6 math problems correctly, for example, has little meaning unless we also know that there were 10 math problems in all. For some behaviors there is no maximum or number. For example, there is no maximum number of times a student may call out or leave her seat during class.

2. *Rate:* The rate of behavior is frequency expressed in a ratio with time.
   Brett got out of his seat 0.2 times per minute.
   Yao did 0.6 math problems per minute *during a 2-minute timed trial.*
   Marvin had 1.3 tantrums per hour *in a 6-hour school day.*
   Lois put her hand in her mouth 0.5 times per minute *during a 10-minute story time.*

If all the observation periods are the same length, one simply reports the number of occurrences and the length of the observation periods. Rate, however, is most often used to compare the occurrence of behavior among observation periods of different lengths. Converting frequency data to rate data enables us to compare data if we are unable to standardize observation periods or opportunities to respond. It makes it possible to compare data, for example, if observation periods are interrupted or if worksheets have different numbers of problems. Rate is calculated by dividing the number of times a behavior occurred by the length of the observation period. For example, if Brett got out of his seat 6 times during the 30-minute math lesson on Monday morning, his rate is 0.2 per minute (6 occurrences divided by 30 minutes). If he left his seat 8 times during a 40-minute social studies class with the fourth grade, his rate is still 0.2 per minute (8/40). The rate is the same across observation periods and, in this example, across settings.

3. *Duration:* The duration of a behavior is a measurement of how long a student engages in it
   Brett was out of his seat for a total of 14 minutes.
   Brett was out of his seat an average of 3 minutes per instance.
   Yao worked on her math for 20 minutes.
   Marvin's tantrum lasted for 65 minutes.
   Lois had her hand in her mouth for 6 minutes.

   Duration is important when the concern is not the number of times Brett gets out of his seat, but how long he is out of his seat each time he gets up or how long he stays up during a given observation period. He may leave his seat only twice during a 40-minute lesson, but if he stays up for several minutes each time, that is a different problem from popping up and going right back down again. If we record the duration of Brett's out-of-seat behavior, we can state that he was out of his seat for a total of 8 minutes during the 30-minute class, or we can report the length of each instance, or we can compute the average amount of time he spent out of his seat during each instance.

4. *Latency:* A behavior's latency is the length of time between instructions to perform it and the occurrence of the behavior.
   After I told Brett to sit in his chair, it took him 50 seconds to sit down.
   After the teacher said, "Get to work," Yao stared into space for 5 minutes before she started her math.
   It took 20 minutes for Marvin to become quiet *after I put him in time-out.*
   After I told Lois to take her hand out of her mouth; it was 2 minutes *before she did so.*

   Latency is relevant when the concern is not how long it takes a student to do something, but how long it takes to begin to do it. For example, Yao may solve 60% of her math problems correctly within an acceptable amount of time once she *starts*, but it takes her 7 minutes to get started.

5. *Topography:* The topography of behavior is the "shape" of the behavior—what it looks like.
   Yao writes all the 4s backwards *on her math paper.*
   Marvin screams, kicks his heels on the floor, and pulls his hair *during a tantrum.*
   Lois's hand sucking involves her putting her fingers in her mouth up to the knuckles.

   Topography describes a behavior's complexity or its motor components. A tantrum, for example, may involve many behaviors performed simultaneously. Some behaviors consist of a chain, or sequence, of individual responses that usually occur together.

6. *Force:* The force of behavior is its intensity.
   Yao writes so heavily that she makes holes in her paper.
   Marvin screams so loudly that the teacher three doors down the hall can hear him.
   Lois's hand sucking is so intense that she has broken the skin on her thumb.

   Describing the intensity or force of a behavior often results in a qualitative measure that is hard to standardize. We are attempting to communicate how loud a scream is (usually without the use of an audiometer), how hard a child is banging a table, or how forcefully he is hitting himself or another child.

7. *Locus:* The locus of a behavior describes where it occurs, either in the environment or, for example, on the child or victim's body.
   Brett walks to the window *and stares outside.*
   Yao writes the answers to her math problems in the wrong spaces.
   Marvin hits his ears *during a tantrum.*
   Lois sucks the fingers of her left hand.

   Locus describes either the target of the behavior or where in the environment the behavior is taking place.

The Professor Effects a Rescue

As Professor Grundy was walking to his car, he observed a congregation of students, including DeWayne, gathered around something that was not visible to him. His curiosity aroused, he strolled over to the group. As he got closer, he observed the object of the students' interest, which was an extremely large white dog. The animal was panting, its head was drooping, and it appeared emaciated. Its coat was matted and filthy, and it was dragging about 3 feet of chain from a metal choke collar drawn tightly around its neck.

"Look, Professor," said DeWayne, "I think it's a white St. Bernard. Do you think it bites?"

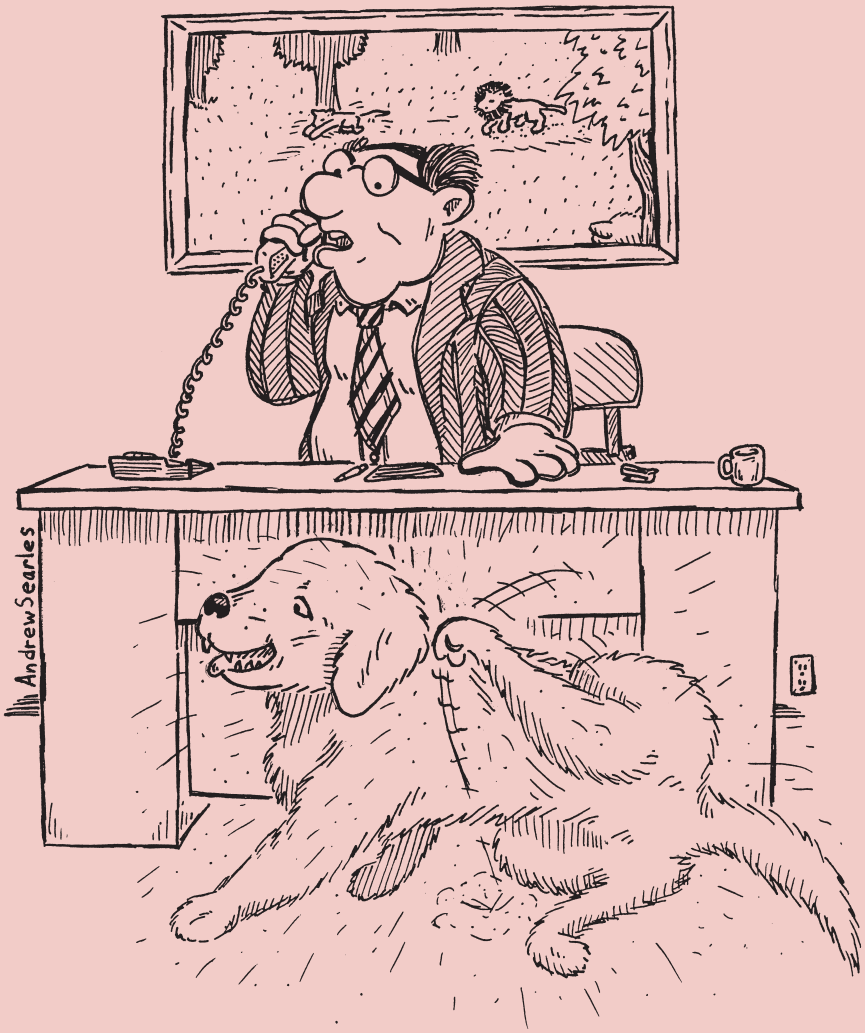
"See here," said the professor firmly, "it is dangerous to approach strange dogs. Someone should call the campus police and tell them to alert the city animal control officers." The dog, apparently determining that the professor was the highest authority present, staggered over to him, rested his huge head against Grundy's leg, and gazed at him soulfully with large, brown eyes.

"On the other hand," said the professor, taking the end of the chain and gently tugging, "perhaps I'll just make a call myself." The professor returned to his department with the dog. As he passed the departmental secretary, she gasped with alarm and began, "Professor, you can't . . ."

Grundy took his stopwatch from his pocket, activated it, and handed it to her. "This won't take 5 minutes," he said. "Time the duration and see for yourself." The professor placed a call to a colleague at the veterinary school who, upon hearing Grundy's description of the animal and its condition, stated, "What you've got there is a Great Pyrenees. What's he doing? I can hardly hear you."

"What he is doing," replied the professor, "is scratching. The topography of the behavior is that he is using his left hind foot to scratch behind his left ear. His foot is moving at a rate of 75 movements per minute according to the second hand on my watch. The force is sufficient to scatter dog hair and various other debris over a 3-foot radius, and his foot is hitting the floor between every third to fourth scratch with sufficient force to be heard in the lobby. He has been scratching for 3 minutes now, and he began scratching within 15 seconds of entering my office.

"Oh," said the veterinarian a little blankly, "probably fleas." (The professor looked furtively down the hall hoping the secretary couldn't hear.) "Why don't you bring him by the clinic and we'll check him out. It sounds like he's been on the road for a while. If he's



healthy, we can see about getting in touch with a rescue society. They have a lot of trouble placing those big guys, though."

As Grundy returned through the lobby, retrieving his stopwatch from the secretary and confirming the duration of his stay as 4 minutes, 34 seconds, the secretary said, "He's really a sweet bunny of a boy, isn't he?" The dog wagged his long, plumy tail weakly. "Look, Professor, I think the sweetie likes me."

"His name," stated the professor firmly, "is Burrhus."

The decision to use a particular system of data collection is based partly on the dimension of behavior that is of concern and partly on convenience. Systems for collecting data can be classified into three general categories. The first is recording and analyzing written reports that ideally include a full record of behaviors emitted during an observation period. The second is the observation of tangible products resulting from a behavior. The third is recording a sample of the behavior as it occurs. These systems may be categorized as follows:

Analyzing written records:	Anecdotal reports
Observing tangible products:	Permanent product recording
Observing a sample of behavior:	Event recording
	Interval recording
	Time sampling
	Duration recording
	Latency recording

ANECDOTAL REPORTS

Anecdotal reports are written to provide as complete a description as possible of a student's behavior in a particular setting or during an instructional period. Anecdotal reports do not identify a predefined or operationalized target behavior. After recording and analyzing data, the observer expects to identify a specific behavior that needs changing. Anecdotal reports are useful primarily for analysis, not for evaluation.

Teachers, parents, and therapists frequently use an anecdotal system of data collection to describe some general disturbance that is taking place or a lack of academic progress. For example, it might be reported that "Sheila constantly disrupts the class and does not complete her own work" or "During therapy sessions, I cannot seem to get Sheila under control to do the needed speech remediation."

Reports such as these are common and should prompt the applied behavior analyst to pinpoint the behavior (see Chapter 2). Should the specific behavior continue to elude identification, the analyst must further isolate and identify a target behavior that may be the source of the complaint in the natural setting of the behaviors—such as at the dinner table or in the classroom during reading period—and attempt to write down everything that occurs.

This system of data collection produces a written description of nearly everything that occurred in a specific time period or setting. It results in a report written in everyday language, describing individuals and interactions, rather than isolated marks on a data sheet. Wright (1960) provided some guidelines for writing anecdotal reports:

Guide to writing anecdotal records.

1. Before beginning to record anecdotal data, write down the setting as you initially see it, the individuals in the setting and their relationships, and the activity occurring as you are about to begin recording (for example, lunch, free play).
2. Include in your description everything the target student(s) says and does and to whom or to what.
3. Include in your description everything said and done to the target student(s) and by whom.
4. As you write, clearly differentiate fact (what is actually occurring) from your impressions or interpretations of cause or reaction.
5. Provide some temporal indications so you can judge the duration of particular responses or interactions.

STRUCTURING AN ANECDOTAL REPORT

After observations have been made, an anecdotal report must be analyzed to determine the behavior(s), if any, that should be the subject of a behavior-change program. The observations in this initial anecdotal format are difficult to separate into individual behaviors and relationships, so it is helpful to present the anecdotal data in a more

FIGURE 3–1 Excerpt from an anecdotal report.

9:40 a.m.: Brian is walking around the room, touching various things such as plants on the windowsill. Teacher says, It 's now time for reading group. Everybody bring your books to the round table. You too, Brian." Teacher goes to table. Brian continues to wander. Teacher, in louder voice, "I'm still waiting." She goes and puts her hand on his shoulder. Brian pulls his shoulder out from under her hand. She takes him by the hand to the group of four other students. Brian sits. Teacher says, "Open your books. Where is your book, Brian?" Brian says, "Back there." "Back where, Brian?" "In my desk." "Go get it." "I will read from her book." "No, Brian, please go get your own book" (about 15 seconds pass). "Now, Brian, we are all waiting for you." Brian says, "So wait, we have plenty of time." Teacher stands. Brian gets up, goes to his desk (where he sits). When first student, Larry, is finished reading, teacher says, "Brian, come back here. It's almost your turn to read." Brian comes back to the table. Carl is reading. Brian makes a noise with his nose. Karen, sitting to his left, giggles and says, "Yuk." Teacher tells Karen to stop talking. Brian makes nose noise again. Karen, "Oh, yuk, yuk." Teacher says, "Brian, I see you. Stop it. Do we all see Brian? That is no way to behave when we are learning." Brian drops his book, bends down to get it, his chair falls. Teacher tells him to "come sit next to me." Brian moves his chair and begins to hum quietly. Teacher stands, moves three feet away from table, and tells Brian to move his chair "over here away from the group." Back at table she says, "Do we all see Brian and what happens when you disturb the group?" Larry raises his hand. "Yes, Larry." Larry says, "They don't get to read." Teacher says, "Yes, very good, Larry. Now let's read again: your turn, Mary." Mary starts reading. Brian is rocking in his chair, Karen looks and giggles. Brian continues to rock: his chair falls backward. Teacher reprimands him, takes him to front of room, and puts him in chair facing blackboard. Brian is singing. Larry yells, "Stop singing. You're disturbing me." Brian quiets down, begins drawing on blackboard (teacher is seated with her back to him). Brian intermittently sings loud enough to be heard. Twice teacher says, "Quiet down, Brian." Reading group ends (17 minutes later). Students are told to line up at the door. On the way out teacher tells Brian how good he was while separated from the group. "But tomorrow you will have to read first in group." 10:35 a.m.: Brian is sent to P.E. with classmates.

schematic manner for review. Bijou, Peterson, and Ault (1968) employed a system for sequence analysis in which they redrafted an anecdotal report into a form that reflects a behavioral view of environmental interactions. By this system, the contents of the report are arranged into columns divided to indicate antecedent stimuli, specific responses, and consequent stimuli. This table format clearly represents the temporal relationship among individual behaviors, the antecedents that stimulate them, and the consequences that maintain them.

The anecdotal report in Figure 3–1 was taken in an elementary classroom. It records a period of interaction between a student named Brian, his teacher, and the members of his reading group.

Using the approach suggested by Bijou and his colleagues, the beginning of this report could be transposed into columns as begun in Figure 3–2. The antecedents, behaviors, and consequences are numbered to indicate the time sequence. Note that transposing the report makes it apparent that in several instances, consequences of a given response can become the antecedents for a succeeding response.

When the content of an anecdotal report has been arranged in a format that clearly presents the sequence of and the relationships among behavioral events, the source of the problem behavior may be determined. The following questions help in analysis:

Questions for analyzing anecdotal information.

1. What are the behaviors that can be described as *inappropriate?* The behavior analyst should be able to justify labeling the behaviors as inappropriate, given the setting and the activity taking place.
2. Is this behavior occurring frequently, or has a unique occurrence been identified?
3. Can reinforcement or punishment of the behavior be identified? Teachers, parents, other children, or some naturally occurring environment event may deliver consequences, intentionally or otherwise.

FIGURE 3–2 ***Structure of an anecdotal report.***

Time	Antecedent	Behavior	Consequence
9:40 a.m.		1. Brian is walking around room.	
	2. Teacher: Time for group. . . . "You too, Brian." T moves to table.		
		3. B continues to walk.	
			4. T: "I'm still waiting."
	5. T puts hand on B's shoulder.		
		6. B pulls shoulder.	
			7. T takes B's hand and leads him to table.
		8. B sits.	
	9. T: "Where is your book, B?"		
		10. B: "Back there."	
			11. T: "Back where?"
		12. B: "In my desk."	
			13. T: "Go get it."
		14. B: "I will read from her book."	

4. Is there a pattern to these consequences?
5. Can antecedents to the behavior(s) be identified?
6. Is there a pattern that can be identified for certain events or stimuli (antecedents) that consistently precede the behavior's occurrence?
7. Are there recurrent chains of certain antecedents, behaviors, and consequences?
8. Given the identified inappropriate behavior(s) of the student and the patterns of antecedents and consequences, what behavior really needs to be modified, and who is engaging in the behavior (for example, the referred student, the teacher, or the parent)?

The use of anecdotal reports is not always practical for general education teachers. Special education teachers may be called upon to observe students who are having behavioral or academic difficulty or who are in the process of being referred for special education services. For such observation, skill in recording and analyzing anecdotal data is extremely valuable. Anecdotal reports can enable these teachers to determine what factors in the classroom are occasioning or maintaining appropriate and inappropriate behaviors. This information will serve as the basis for making decisions about possible changes in the classroom environment or in behavior-management strategies. Anecdotal observation may also be used as a first step in a longer process for dealing with persistent, highly disruptive, or seriously harmful behaviors. This process, known as functional assessment (see Chapter 6), requires detailed observation, analysis, and manipulation of objects and events in a student's environment to determine what is occasioning and maintaining the behaviors.

PERMANENT PRODUCT RECORDING

Permanent product recording is the easiest to use, but not all behaviors leave a permanent product.

Teachers have been using **permanent product recording** since the first time a teacher walked into a classroom. A teacher uses permanent product recording to grade a spelling test, verify the creation of a chemical emulsion, or count the number of cans a student has placed on a shelf. *Permanent products* are tangible items or environmental effects that result from a behavior. Permanent products are outcomes of behavior; thus, this method is sometimes called *outcome recording*. This type of recording is an *ex post facto* method of data collection.

To collect permanent product data, the teacher reviews the statement of the behavior as written in the behavioral objective and determines what constitutes an acceptable outcome of the behavior. For example, if the behavior is building a tower of blocks, the objective states whether the student is required to place one block on top of another or whether the blocks should be arranged in a certain color sequence. If the behavior is academic, conditions also are specified. For example, the objective may specify the number of spelling errors permitted in a written paragraph or the number of references required in a term paper. If the behavior is vocational, quality may be specified as well as the number of widgets to be assembled. In each case, the teacher reviews the operational definition of the behavior. After evaluating the products of the required behavior, the teacher simply notes how many of the products were produced and how many were acceptable according to the definition.

Because the concrete results of a behavior are being evaluated and recorded, the teacher does not have to observe the student directly engaged in the behavior. Convenience is the explanation for the frequent use of permanent product recording in the classroom: It causes minimal interference with a classroom schedule.

Applications of permanent product recording.

The versatility of permanent product recording makes it useful in a variety of instructional programs and settings. In the home, permanent product recording has been used to record data on incontinence (Adkins & Matthews, 1997) and room cleaning (Dyer, Schwartz, & Luce, 1984). In vocational settings this method has been used to record outcomes of tasks ranging from pot scrubbing (Grossi & Heward, 1998) to stuffing envelopes, sorting laundry, and assembling nuts and bolts (Worsdell, Iwata, & Wallace, 2002). In educational settings it has been used to record data on basics such as cup stacking (Zarcone, Fisher, & Piazza, 1996), milk drinking (Tiger & Hanley, 2006), and cursive writing (Trap, Milner-Davis, Joseph, & Cooper, 1978); and diverse academic tasks such as spelling accuracy (Coleman-Martin & Heller, 2004; Grskovic & Belfiore, 1996) and number of words typed by students with physical disabilities using word prediction software (Tumlin & Heller, 2004; Mezei & Heller, 2004); number of words written in classroom journal entries and rubric scoring of quality (e.g., grammar, sentence structure, and organization) (Regan, Mastropieri, & Scruggs, 2005); number of words in story writing and number of story elements (e.g., main characters, locale, time, actions) (Lienemann, Graham, Leader-Janssen & Reid, 2006; Reid & Lieneman, 2006), and quality of paragraphs (e.g., topic sentence, transition words) (Konrad, Trela, & Test (2006); completion and accuracy of arithmetic problems (Freeland & Noell, 1999; Jolivette, Wehby, Canale, & Massey, 2001; Neef, Nelles, Iwata, & Page, 2003; Wood, Frank, & Wacker, 1998); and performance on science and foreign language quizzes (Cavanaugh, Heward, & Donelson, 1996; Lloyd, Eberhardt, & Drake, 1996). Permanent product recording was also the method used to assess reduction of bathroom graffiti (Mueller, Moore, Doggett, & Tingstrom, 2000), note taking during lectures (Neef, McCord, & Ferreri, 2006), and to monitor completion and accuracy of homework assignments (Hinton & Kern, 1999; Ryan & Hemmes, 2005).

The main advantage of permanent product recording is the durability of the sample of behavior obtained. The permanent product is not apt to disappear before its occurrence can be recorded. In light of this, the teacher may keep an accurate file of the actual

products of certain target behaviors (such as test papers) or a report of the products for further review or verification later.

Permanent product recording may include the use of audiotape, videotape, and digital recording systems. With recording equipment, teachers can make samples of specific transitory behaviors that would not ordinarily produce a permanent product. Samples of behavior in hectic settings such as play groups can be recorded and analyzed at leisure. Samples of behaviors from nonschool settings such as a student's home can be made by parents and brought to professionals for analysis. Individual and group samples of expressive language, for example, have been audiotaped (Matson, Sevin, Fridley, & Love, 1990; Orsborn, Patrick, Dixon, & Moore, 1995) and videotaped (Kim & Hupp, 2007; Loncola & Craig-Unkefer, 2005; Schlosser, Walker, & Sigafoos, 2006). Samples of student performance in general and special education settings have been videotaped to allow collaboration among interdisciplinary team members to determine educational goals and intervention techniques (Anderson, Hawkins, Hamilton, & Hampton, 1999). Audiotaping and videotaping allow for data collection after the fact, just as grading a student's exam or composition after school does.

Behavioral dimensions for which permanent product recording may be used.

What permanent products or outcomes might be observed for each of the behavioral dimensions discussed in the section on choosing a system of data collection?

Rate: number of written products of any academic behavior per unit of time

Duration or **latency:** unfortunately do not lend themselves to permanent product recording unless recording equipment is available

Topography: the correct formation of letters or numerals; following a pattern in such activities as pegboard designs, block building, or vocational assemblies

Force: too light, too heavy, or uneven pressure when writing or typing; holes kicked in a classroom wall by a student having a tantrum

This list of examples is by no means exhaustive. Because permanent product recording is relatively simple and convenient, teachers can be imaginative in defining behaviors in terms of their outcomes. We have known teachers who operationally defined

- *test anxiety* as the number of visible erasures on a test paper
- *sloppiness* as the number of pieces of scrap paper on the floor within 2 feet of a student's desk
- *hyperactivity* as the number of table tennis balls still balanced in the pencil tray of a student's desk

The following vignette examines one use of permanent product recording.

Mr. Martin Observes Room Cleaning

Mr. Martin, while majoring in special education, was a night-shift assistant at a residential treatment facility for students with severe emotional and behavioral problems. One of his duties was to see that each bedroom was cleaned before bedtime. He decided to establish some system for reinforcing room cleaning but was uncertain about what he should measure. When he tried measuring and reinforcing the time students spent cleaning their rooms, he found that although there was a great deal of scurrying around, the rooms were still very messy. Because the major problem appeared to be clothes, toys, and trash scattered on the floors, beds, and other furniture, he decided to use the number of such objects as his measure. Each evening before lights-out, he stood at the door of each bedroom with a clipboard containing a sheet of paper with each resident's name and a space for each day of the week. He rapidly counted the number of separate objects scattered in inappropriate places and entered the total in the space on his data sheet.

Observational Recording Systems

Whereas the permanent product method of data collection records the outcome of a behavior, **observational recording systems** are used to record behavior samples as the behavior is actually occurring. A data collector may choose from several basic observational recording systems. Teachers who are interested in recording the number of times a behavior is occurring may select **event recording**. Those who want to find the proportion of a specified time period during which the behavior occurs may select **interval recording** or **time sampling. Duration recording** allows the teacher to determine the length of time the student spends performing some behavior. **Latency recording** measures the length of time it takes a student to start doing something. An illustration of the relationship between observational recording procedures and the components of a behavioral, stimulus-response sequence is shown in Figure 3–3.

Event recording provides an exact count of how many times a behavior occurs.

Event Recording

Event recording is a frequently used observational recording procedure because it most directly and accurately reflects the number of times a behavior occurs. When using event recording, the observer makes a notation every time the student engages in the target behavior. Tallying these notations gives an exact record of how often the behavior occurred. A count of the target behavior is made during a specified observation period—for example, during a reading period or during lunch. Recording how often the behavior occurs within a given time period documents its frequency. If the lengths of the observation periods are constant, the observer may simply report the number of times the behavior occurred, its frequency, or its rate, the number of times it occurred per minute or hour during that period. Rate may also be reported if the observation periods vary in length. Another strategy is arbitrarily to standardize the length of the observation period—for example, by taking data for just the first 20 minutes of the period each day.

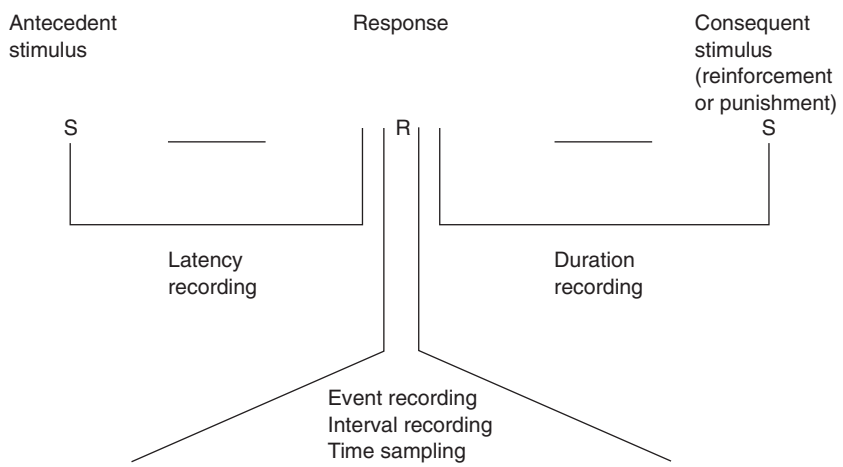


FIGURE 3–3 ***Observational data collection systems as related to the basic behavioral paradigm.***

Event recording can be used only for discrete behaviors.

Event recording is usually the method of choice when the objective is to increase or decrease the number of times a student engages in a certain behavior. Event recording may be used to record an increase in an appropriate social behavior, such as counting the number of times a student shares a toy with a classmate. It may be used to record an increase in an academic response (counting the number of correctly defined science vocabulary words) or a decrease in an inappropriate behavior (counting the number of times a student curses during physical education class). Because the teacher attempts to record the exact number of times the behavior occurs, event recording must be used behaviors that are discrete. **Discrete behaviors** have an obvious, or agreed upon, beginning and end. The observer can make an accurate frequency count because she can clearly judge when one occurrence ends and the next begins. Event recording has been used for counting and recording behaviors in a range of content areas, including *academics:* picture naming (Stromer, MacKay, McVay, & Fowler, 1998), words spelled with a voice-output communication device (Schlosser, Blischak, Belfiore, Bartley, & Barnett, 1998), word recognition and oral reading (Alber-Morgan, Ramp, Anderson, & Martin, 2007; Allen-DeBoer, Malmgren, & Glass, 2006; Browder & Minarovic, 2000; Eckert, Ardoin, Daly, & Martens, 2002; Hurst & Jolivette, 2006); addition facts (Jolivette, Lingo, Houchins, Barton-Arwood, & Shippen, 2006), money selection and counting (Denny & Test, 1995; Schloss, Kobza, & Alper, 1997), and science vocabulary words (Johnson, Schuster, & Bell, 1996); *communication:* yes/no responses (Neef, Walters, & Engel, 1984), question asking (Williams, Perez-Gonzalez, & Vogt, 2003), requesting (Schlosser et al., 2006), and manual signs (Partington, Sundberg, Newhouse, & Spengler, 1994); *self-help:* accepting bites of food (Piazza, Patel, Gulotta, Seven, & Layer, 2003), drinking from a cup (Hagopian, Farrell, & Amari, 1996), and urinary control (Simon & Thompson, 2006); *social skills:* waiting one's turn (Laushey & Heflin, 2000), social initiations (Nelson, McDonnell, Johnston, Crompton & Nelson, 2007), and saying "please" (Drasgow, Halle, & Ostrosky, 1998); *leisure skills:* soccer (Brobst & Ward, 2002), football (Smith & Ward, 2006), basketball (Vollmer & Bourret, 2000), and toy activation using a switch (Mechling, 2006). Event recording has also been used to count instances of *inappropriate behavior*, such as talk-outs (Crozier & Tincani, 2005), eye poking (Smith, Russo, & Le, 1999), hand mouthing (Lerman & Iwata, 1996), hand flapping (Conroy, Asmus, Sellers, & Ladwig, 2005), face slapping, head banging (Sigafoos, Penned, & Versluis, 1996), throwing, hitting, spitting, kicking, slapping and pushing (Lien-Thorne & Kamps, 2005; Singh, Lancioni, Winton, et al., 2006), eating cigarette butts (McCord, Grosser, Iwata & Powers, 2005), and dropping to the floor (Smith & Churchill, 2002).

Event recording may also be used when teaching from a task analysis. A task analysis is a list of the individual steps that when chained together form a complex behavior such as hand washing or solving addition problems. During instruction the teacher records the performance of the steps listed in the task analysis. The steps of a task analysis are a series of discrete behaviors. The teacher can thus record a student's performance on each step as an individual, discrete behavior. A task analysis for teaching a secondary-age student with moderate disabilities to clean tables, for example, might consist of 12 steps, the first 5 of which could be: (1) place bucket on table, (2) pick up cloth, (3) place cloth in soap and water, (4) wring out cloth, and (5) rub section one of table (Smith, Collins, Schuster, & Kleinert, 1999). Each of these steps is a discrete behavior with a clear, or clearly definable, beginning and ending. Progress toward acquiring individual steps or the entire task of table cleaning can be recorded through event recording (see Figures 9–3 and 9–4 for a sample data sheet).

Examples of event recording with task analysis include skills in the school such as using a calculator, operating an audio tape recorder, accessing a computer program, or sharpening pencils (Werts, Caldwell, & Wolery, 1996), and playing red light/green light and Simon Says (Arntzen, Halstadtro, & Halstadtro, 2003); in the home for self-help skills such as dressing (Sewell, Collins, Hemmeter, & Schuster, 1998), self-catheterization (McComas, Lalli, & Benavides, 1999), and self-administration of insulin and home care such as making coffee, washing dishes, or washing windows (Steed & Lutzker, 1999); in the community for skills such as cashing a check, crossing a street, mailing a letter (Branham, Collins, Schuster, & Kleinert, 1999), using an ATM and making a purchase with a cash card (Cihak, Alberto, Taber-Doughty & Gama, 2006), and counting money using the one-more-than method (Denny & Test, 1995). A complete discussion of task analysis appears in Chapter 9.

Some labels for behavior may be used to describe what are actually a number of different responses, each of which may or may not occur each time the so-called behavior occurs. Examples that many teachers try to target for change are on-task or off-task behavior, appropriate or inappropriate verbalizations, in-seat or out-of-seat behavior, or disturbing one's neighbor. For accurate event recording, a standard definition, with an agreed-upon beginning and end, is necessary. In other words, such behaviors can be made discrete by defining them operationally (see Chapter 2).

Certain behaviors, however, may not be adequately measured using event recording. This data collection procedure is not appropriate in the following instances:

1. Behavior occurring at such a high **frequency** that the number recorded may not reflect an accurate count. Certain behaviors, such as the number of steps taken while running, some stereotypic behaviors (such as hand flapping or rocking by students with severe disabilities), and eye blinking may occur at such high frequencies that it is impossible to count them accurately.
2. Cases in which one behavior or response can occur for extended time periods. Examples of such behaviors might be thumb sucking or attention to task. If out-of-seat behavior were being recorded, for instance, a record showing that the student was out of her seat only one time during a morning would give an inaccurate indication of what the student was actually doing if this one instance of the behavior lasted from roll call to lunchtime.

An advantage of event recording, in addition to accuracy, is the relative ease of the data collection itself. The teacher does not need to interrupt the lesson to take data. The teacher may simply make a notation on an index card or paper on a clipboard, make slash marks on a piece of tape around his wrist, or transfer paper clips from one pocket to another. This information may be tallied and transferred to a data sheet similar to the one presented Figure 3–4.

Event recording is also easily used for many academic behaviors. Figure 3–5 is a data sheet used for recording the errors made during an oral reading exercise. The teacher simply places a mark in the appropriate row as a particular error is made. The column heads may record the day of week, the child who is reading, dates, page number from the reader where the mistake was made, and so on.

Using the data sheet in Figure 3–6, a teacher can record the correct verbal reading of sight words. The sight words chosen for the student are listed in the left column.

The succeeding columns provide space to indicate whether the student's reading of the word was correct or incorrect. The total number (or percent) correct is recorded at the bottom of the column (on April 12th, for example, Deepa read 5 words correctly—though not the same 5 as on April 10th). Data on the number of words identified correctly may

Event Recording Data Sheet

Student: PATRICIA
Observer: MRS. COHEN

Behavior: INAPPROPRIATE TALK-OUTS (NO HAND RAISED)

	Time Start Stop	Notations of occurrences	Total occurrences
5/1/95	10:00 10:20	~~////~~ ~~////~~ //	12
5/2/95	10:00 10:20	~~////~~ ////	9

FIGURE 3–4
Basic data sheet for event recording.

Student: JEREMY
Observer: MS. GARWOOD

Behavior: Errors in oral reading

		CAROL	5/1/95	EXERC. #2	READER PG. 7.
SUBSTITUTIONS					
MISPRONOUNCIATIONS					
INSERTIONS					
REPETITIONS					

FIGURE 3–5
Data sheet for event recording.

be taken during instruction as the teacher records Deepa's response as correct or incorrect each time she is asked to read a word. This is referred to as trial-by-trial data collection. Data may be taken instead at the conclusion of the instructional session by giving Deepa the opportunity to read each word one more time and recording only those responses. This is one form of probe data collection. A simple method for recording correct and incorrect answers when using flash cards is to mark directly on the back of the card on which the word (or other cue) is printed. These marks can be transferred later to a summary data sheet.

For the more mechanically inclined, counting devices are commercially available. Although these make data collection easier and more accurate, they entail some expense and may break. An inexpensive counter sold for tallying purchases in a grocery store or golf strokes may be useful. Stitch counters designed to fit on the end of a knitting needle

FIGURE 3–6 ***Data sheet for event recording.***

Student: **Inclusive Dates:**

OBJECTIVE: Given the following list of 12 sight words found in general merchandise stores (e.g., Wal-Mart), Deepa will verbally state each.

Criterion: with 100% accuracy for 4 consecutive sessions.

Dates/Trials

Items	4/10	4/12								**Comments**
restrooms	+	+								
exit	+	—								
girls	—	+								
housewares	—	—								
pet	+	—								
supplies	—	+								
checkout	—	—								
express	—	+								
shoes	+	—								
linens	—	+								
videos	+	—								
electronics	—	—								
Total Correct	5/12	5/12								

come in sizes large enough to fit on a pen. Drash, Ray, and Tudor (1989, p. 453) describe the following four-step procedure for using most simple, inexpensive pocket calculators to record event data:

1. Press number 1
2. Press 1 key
3. Press 5 key
4. Press 5 key to record each subsequent event

These steps set up the calculator so that each time the 5 key is pressed, it will cumulatively add each occurrence of the behavior. At the end of the observation period, the

observer subtracts 1 from the cumulative total (because 1 was recorded during the initial setup), and the exact total of occurrences for the observation period will appear on the calculator. With practice, and perhaps a small piece of tape on the 5 key as a prompt, an observer can do this with the calculator in her pocket. (Anyone who attempts to teach wearing garments without pockets needs a carpenter's apron.)

Recording Controlled Presentations

One variation on the event-recording technique is the use of **controlled presentations** (Ayllon & Milan, 1979). In this method, the teacher structures or controls the number of opportunities the student will have to perform the behavior. Most often this method consists of presenting a predetermined number of opportunities, or trials, in each instructional session. A **trial** may be viewed as a discrete occurrence because it has an identifiable beginning and ending. A trial is defined by its three behavioral components: an antecedent stimulus, a response, and a consequent stimulus (S-R-S). The delivery of the antecedent stimulus (usually a verbal cue) marks the beginning of the trial, and the delivery of the consequent stimulus (reinforcement, correction, or punishment) signifies the termination of the trial. For example, in a given session the teacher may decide that a student will be given 10 opportunities, or trials, to respond by pointing to specified objects upon request. Each trial is then recorded as correct or incorrect. Controlled presentation allows the teacher to monitor progress simply by looking at the number of correct responses for each session.

Figures 3–7 and 3–8 present variations of data sheets used for the collection of discrete trial or controlled presentation data. The data sheet in Figure 3–7 (variation of Saunders & Koplik, 1975) is arranged from left to right for 15 sessions. Within each session, or column, there are numbers representing up to 20 trials. The teacher records dichotomous data (whether the response was correct or incorrect) using the following simple procedure:

After each trial:

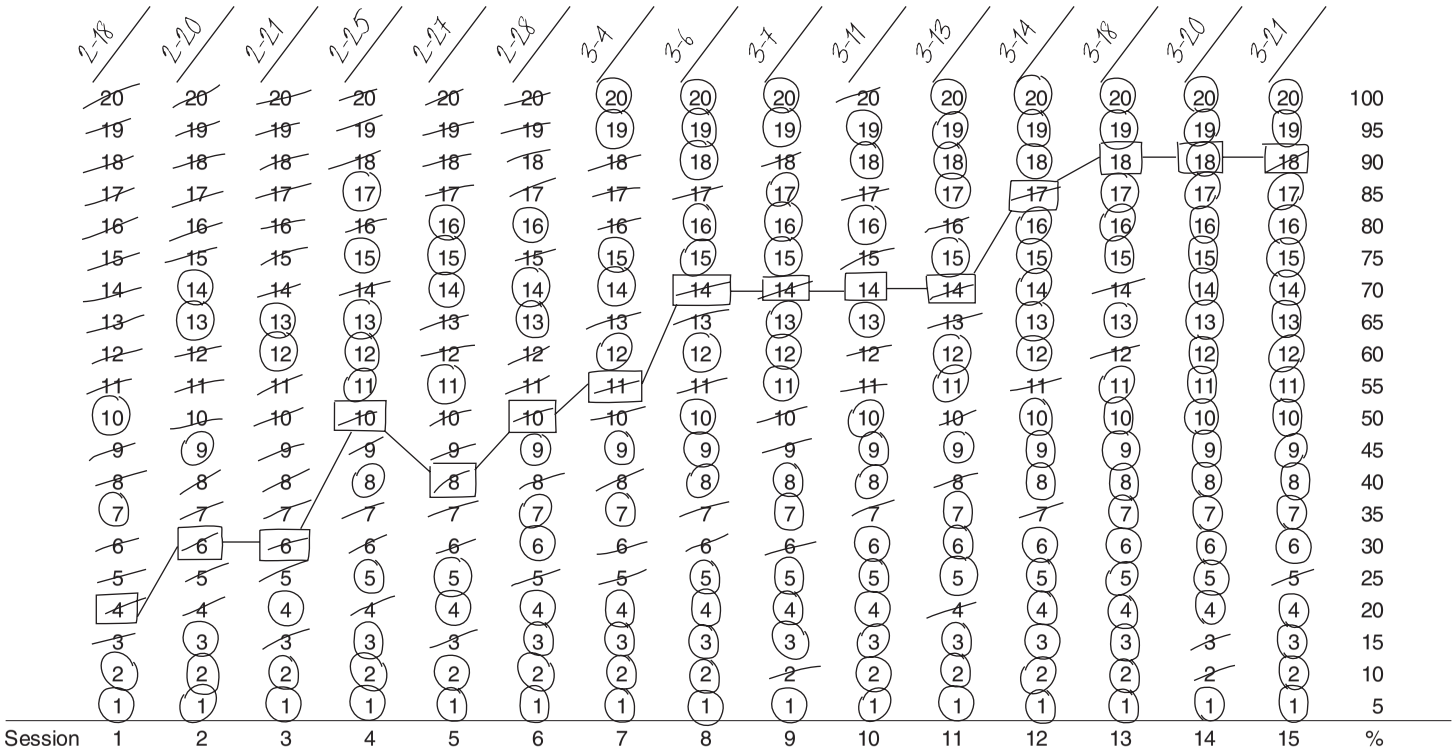
1. Circle the trial number that corresponds to a correct response.
2. Slash (/) through the trial number that corresponds to an incorrect response.

After each session:

1. Total the number of correct trials (those circled).
2. Place a square around the corresponding number in the session column that corresponds to the number of correct trials.
3. To graph directly on the data sheet, connect the squared numbers across the sessions to yield a learning curve.
4. The column on the far right allows the number of correct trials per session (the number with the square around it) to be converted to the percentage of trials correct. If the number of correct trials in a 20-trial session was 8, looking at the last column, we see that the percentage correct is 40.

Figure 3–8 is a modification of the previous data sheet that allows the observer to use it for up to six students working on the same task or for up to six tasks for the same student. Controlled presentations and specifically constructed data sheets are also used for two additional instructional needs. Data sheets may be constructed to use when teaching chained tasks through the use of a task analysis, and for use when the teacher wants to record prompts (e.g., verbal or physical assistance) used to assist student performance during instruction. These instructional and data-recording strategies are discussed in Chapter 9.

FIGURE 3–7 ***Data collection sheet for use with controlled presentations.***



Target behavior/skill Sorting Full vs Not Full

Criterion 90% correct trials for 3 consecutive sessions

Materials restaurant containers for salt, pepper, sugar, ketcup, mustard, napkins

Student Carmen

Note: Adapted from "A Multi-Purpose Data Sheet for Recording and Graphing in the classroom," by R. Saunders and K. Koplik, 1975, *AAESPH Review*. Copyright 1975 by the Association for the Severely Handicapped. Reprinted with permission.

FIGURE 3–8 ***Data collection sheet for use with controlled presentations.***

Name: Peter
Task: Functional Counting to 10
Date:

10	10	10	10	10
9	9	9	9	9
8	8	8	8	8
7	7	7	7	7
6	6	6	6	6
5	5	5	5	5
4	4	4	4	4
3	3	3	3	3
2	2	2	2	2
1	1	1	1	1

Comments:

Name: Calvin
Task: Functional Counting to 10

10	10	10	10	10
9	9	9	9	9
8	8	8	8	8
7	7	7	7	7
6	6	6	6	6
5	5	5	5	5
4	4	4	4	4
3	3	3	3	3
2	2	2	2	2
1	1	1	1	1

Comments:

Name: Tonya
Task: Functional Counting to 10

10	10	10	10	10
9	9	9	9	9
8	8	8	8	8
7	7	7	7	7
6	6	6	6	6
5	5	5	5	5
4	4	4	4	4
3	3	3	3	3
2	2	2	2	2
1	1	1	1	1

Comments:

Name: Rita
Task: Functional Counting to 10
Date:

10	10	10	10	10
9	9	9	9	9
8	8	8	8	8
7	7	7	7	7
6	6	6	6	6
5	5	5	5	5
4	4	4	4	4
3	3	3	3	3
2	2	2	2	2
1	1	1	1	1

Comments:

Name: Dylana
Task: Functional Counting to 10

10	10	10	10	10
9	9	9	9	9
8	8	8	8	8
7	7	7	7	7
6	6	6	6	6
5	5	5	5	5
4	4	4	4	4
3	3	3	3	3
2	2	2	2	2
1	1	1	1	1

Comments:

Name: Roy
Task: Functional Counting to 10

10	10	10	10	10
9	9	9	9	9
8	8	8	8	8
7	7	7	7	7
6	6	6	6	6
5	5	5	5	5
4	4	4	4	4
3	3	3	3	3
2	2	2	2	2
1	1	1	1	1

Comments:

Note: From *Vocational Habilitation of Severely Retarded Adults: A Direct Service Technology*, by G. Bellamy, R. Horner, & D. Inman, 1979, (Baltimore: University Park Press. Copyright © 1979 University Park Press, Baltimore. Reprinted by permission.

Behavioral dimensions appropriate to event recording.

Classroom teachers can improve instruction by using controlled presentations. For example, a teacher might want to be sure to ask each member of a seminar group five questions during a discussion on early Cold War events and the Berlin Wall. A very simple data sheet with the names of the students and a space to mark whether answers were correct or incorrect would provide valuable information for analysis and evaluation.

Event recording (including controlled presentation) lends itself to the observation of the rate of frequency of behavior, for example:

- number of times Mel talks out in an hour
- number of times Charlie hits another student in a 20-minute recess
- number of questions Melissa answers correctly during a 15-minute world geography review
- number of times Sam answers questions in a whisper
- number of times Mary throws trash on the floor
- number of stairs Eliot climbs putting only one foot on each step

Ms. Stallings Counts Tattling

Four of the students in Ms. Stallings's third-grade class seemed to spend most of their time telling her what other students were doing wrong. Ms. Stallings was worried about this for two reasons: the students were not working efficiently and they were driving her up the wall. When she asked her colleague Ms. Barbe for advice, Ms. Barbe suggested that the first thing to do was to find out how often each of the students was tattling.

"Otherwise," she said, "you won't know for sure whether whatever you do to stop them is working."

Ms. Stallings decided to count an instance of tattling each time a student mentioned another student's name to her and described any number of inappropriate behaviors. Thus, "Johnny's not doing his work and he's bothering me," was counted as one instance, but "Harold and Manolo are talking," was counted as two instances. She then went back to Ms. Barbe.

"How can I write it down every time they do it?" she asked. "I move around my room all the time, and I don't want to carry paper and pencil."

Ms. Barbe laughed. "No problem," she answered, "I'm sure that's why dry beans come in so many sizes and shapes. I just pick a different bean for each student, put a handful of each in my right pocket, and transfer by feel to my left pocket when I observe the behavior. Just be sure you get them out before your clothes go into the washer."

Interval Recording and Time Sampling

The observer counts intervals, not discrete behaviors.

Interval recording and time sampling data collection systems are ways of recording an approximation of the actual number of times behavior occurs. Instead of counting each occurrence of the behavior, the teacher counts the number of intervals of time within observation periods during which the behavior occurs. With these methods it is possible to record continuous behaviors (behaviors of longer duration) and high-frequency behaviors that may be incompatible with event recording. Interval recording and time sampling have been used with behaviors such as task engagement (Bryan & Gast, 2000; Massey & Wheeler, 2000; Todd, Horner, & Sugai, 1999), cooperative toy play (Van Camp et al., 2000), and sharing (Reinecke, Newman, & Meinberg, 1999); tantrums and aggressive behaviors such as hitting, kicking, biting scratching, and throwing objects (Charlop-Christy & Haymes, 1998; Lalli, Kates, & Casey, 1999); yelling and bizarre vocalizations (Magee & Ellis, 2000; Wilder, Masuda, O'Connor, & Baham, 2001); stereotypy (Kennedy, Meyer, Knowles, & Shukla, 2000) and self-injury (Irvin et al., 1996; Wacker et al., 1996); and drooling (Lancioni, Brouwer, & Coninx, 1992), hand mouthing (Irvin, Thompson, Turner, & Williams, 1998), and thumb sucking (Friman, 2000).

In terms of making the closest representation of the actual occurrence of the behavior, event recording is the most accurate, interval recording is next, and time sampling is

the least exact (Repp, Roberts, Slack, Repp, & Berkler, 1976). Each system, however, has its advantages and disadvantages.

Interval Recording

When using interval recording, the teacher defines a specific time period (usually between 10 minutes and an hour) during which the target behavior will be observed. This observation period is then divided into equal intervals. These intervals are typically 5, 10, or 15 seconds long, occasionally up to 30 seconds. The shorter the interval is, the more accurate the data. The teacher draws a series of boxes representing the time intervals. In each box or interval the teacher simply notes whether the behavior occurred (+) or did not occur (−) *at any time during the interval.* Therefore, each interval has only one notation. The data sheet for a 5-minute observation period shown in Figure 3–9 has been divided into 10-second intervals. During the first minute of the observation period, the target behavior occurred during two of the intervals, the second and third. Over the total 5-minute period, the target behavior occurred during 12, or 40%, of the intervals.

Interval recording does not provide an exact count of behaviors but is especially appropriate for continuous behaviors.

Because of the way these interval data are recorded, only limited conclusions can be drawn from the record of the behavior's occurrence. Regardless of whether the behavior occurred once or five times during the interval, a single notation is made. Therefore, the actual number of occurrences is not included in the record. If, in the preceding example, cursing was being recorded, all the teacher could say was that, during two of the intervals, the student cursed. There were at least two instances of this behavior, but there may have been more. Even if the student cursed 11 times during the second interval, only one notation would have been made. Recording occurrences of discrete behaviors, such as cursing or hitting, is known as *partial-interval recording* (the behavior does not consume the entire interval).

FIGURE 3–9
Interval recording data sheet.

Student: Darius
Date: 8-29
Time Start: 9:10
Observer: Mrs. Heflin
Behavior: On-task (eyes on paper or writing on paper)
Setting: 4th period math
Time End: 9:15

Length of Intervals in Seconds

Length of Observation Period in Minutes	10"	20"	30"	40"	50"	60"
1'	−	+	+	−	−	−
2'	+	+	−	−	−	+
3'	+	−	−	+	+	+
4'	−	+	+	+	−	−
5'	−	−	−	−	−	−

Number (Percent) of Intervals of Occurrence: 12 intervals, 40% of intervals
Number (Percent) of Intervals of Nonoccurrence: 18 intervals, 60% of intervals

Behaviors such as walking around the room or being off task may begin in one interval and continue unbroken into the next interval. Such timing would appear as two instances, because it would be recorded in two intervals in this instance, but the same duration of behavior would appear as only one instance if it fell within a single interval. Recording ongoing behaviors that may continue for several intervals is known as *whole-interval recording* (the behavior consumes the entire interval).

An additional problem encountered with interval data collection is created by the shortness of each interval in which the notation is to take place. It is very difficult to teach and collect interval data simultaneously. The teacher must keep an eye on a student or students, observe a stopwatch or second hand on a watch, and note the occurrence or nonoccurrence of the target behavior all within a matter of seconds; a third-party observer is often required.

The necessity of looking down at the data sheet to make a recording might cause even an observer to miss an occurrence of the behavior, resulting in inaccurate data. The need to look at one's watch to check the passage of the interval is eliminated by videotaping the observation period and using the time on the VCR to note intervals (Miltenberger, Rapp, & Long, 1999). Observers can also time intervals using an audiotape with beeps at the end of each interval, or set a timer with a seconds indicator and an audible signal, or use commercial products such as the WatchMinder or Invisible Clock that will chime or vibrate at selected intervals. Another way of simplifying the task is to build in opportunities for recording as part of the schedule. For example, the observer might alternate 10-second intervals for observing with 5-second intervals for scoring.

Figure 3–10 is an example of an interval recording sheet for a 15-minute period divided into 10-second intervals. Looking at the notations of occurrence and nonoccurrence, the data collector can infer certain information:

1. approximate number of occurrences of the behavior
2. approximate duration of the behavior within the observation period
3. distribution of the behavior across the observation period

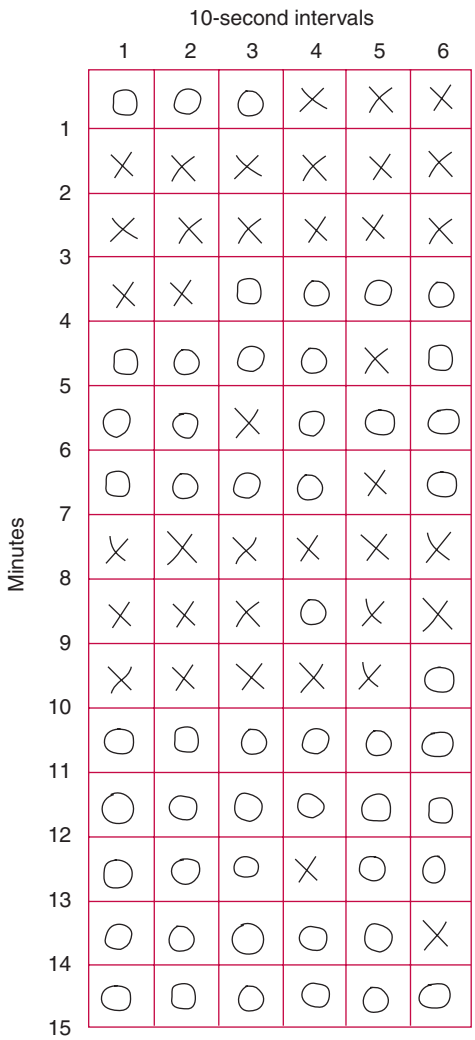
Assuming that off-task behavior during a written arithmetic assignment is recorded in the example, the behavior appears to have occurred in 38 of the 90 intervals. The successive intervals in which the behavior occurred indicate that the off-task behavior occurred over long durations (3 minutes each), but it appears to have been confined primarily to two periods. When reviewing such data, teachers should analyze the situation for some indication of what seem to be the immediate precipitating factors. In this example, off-task behavior may have been due to the two sets of written instructions on the worksheet, which prompted the student to ask a neighbor what to do. Alternatively, if occurrence of the behavior is distributed throughout the time period, apparently lacking in some pattern(s), the teacher may consider whether the off-task behavior is a result of a more generalized reason, such as distractibility due to movement and conversation in the room, boredom with the task, or a lack of prior teacher resulting in a lack of the skills necessary for the task.

Time Sampling

Time sampling allows for only one observation per interval.

In order to use time sampling, the data collector selects a time period in which to observe the behavior and divides this period into equal intervals. This process is similar to that employed with interval recording, but the intervals for time sampling are usually minutes rather than seconds. Such a format allows for observing the behavior over longer periods of time (Ayllon & Michael, 1959). To record these data, the observer draws a series of boxes representing the intervals. The observer simply notes in each box (interval) whether the behavior was occurring (X) or not (O) when the student was observed at the

FIGURE 3–10 ***Interval recording of off-task behavior.***



Student MALCOLM
Date 2/24
Observer MR. RILEY
Time Start 9:15
Time End 9:30
Behavior OFF-TASK

Note occurrence within the 10-second interval.

X = occurrence
O = nonoccurrence

Data summary

Number of intervals of occurrence: 38
Percent of intervals of occurrence: 42%
Number of intervals of nonoccurrence: 52
Percent of intervals of nonoccurrence: 58%

end of the interval. Each interval therefore has only one notation. Note that the time sampling procedure differs from interval recording in that the student is observed only *at the end of the interval* rather than throughout it.

Figure 3–11 shows a data sheet for a 1-hour observation period occurring between 9:05 a.m. and 10:05 a.m. 3 days a week. The hour is divided into six 10-minute intervals. On Monday the target behavior (walking around the room without permission) was occurring at the end of four of the intervals: the first, second, fourth, and fifth. On Wednesday it was occurring at the end of three intervals: the first, fifth, and sixth. On Friday it was occurring at the end of four intervals: the first, fourth, fifth, and sixth. The teacher may summarize the number of intervals at the end of which the behavior was occurring for each day or each week or record the daily average for the week. As time sampling allows for longer periods of observation, the teacher could take data on the student's target behavior for the entire morning. The data sheet in Figure 3–12 allows for a morning's observation period of three hours, with each hour divided into four 15-minute intervals.

FIGURE 3–11 ***Time sampling data sheet for a 1-hour observation period, with 10-minute intervals, 3 days during a week.***

Student: Shane

Behavior: walking around the room without permission

Date: 1/24 1/26 1/28

Time start: 9:05
Time end: 10:05
Total time period per day: 1 hr.

Observer: _____

CODE: X = occurrence O = nonoccurrence (note at the end of each interval)

	10'	20'	30'	40'	50'	60'
1/24	+	+	O	+	+	O
1/26	X	O	O	O	X	X
1/28	X	O	O	X	X	X

Data Summary

Daily: Mon. number of intervals of occurrence:
percent of intervals of occurrence:
number of intervals of nonoccurrence:
percent of intervals of nonoccurrence:

Weekly: number of intervals of occurrence:
percent of intervals of occurrence:
number of intervals of nonoccurrence:
percent of intervals of nonoccurrence:

Average daily: number of intervals of occurrence:
percent of intervals of occurrence:
number of intervals of nonoccurrence:
percent of intervals of nonoccurrence:

Two fairly simple ways of recording time sampling data are to set a timer to ring at the end of the interval and to observe the behavior when the timer rings or to use a tape with prerecorded sounds at chosen intervals that can be heard using an earpiece. To prevent students' figuring out the schedule and performing (or not performing) some behavior only at the end of the interval, the intervals may be of varying lengths. For example, a 10-minute time sampling recording system might have intervals of 8, 12, 6, 14, 9, and 11 minutes. The average interval duration would be 10 minutes, but the students would not know when they were about to be observed. Common sense indicates the need to hide the face of the timer.

An observer may miss a lot of behavior when using time sampling.

Because of the method of recording time sampling data, only limited conclusions can be drawn about the behavior recorded. As with interval recording, the behavior may have occurred more than once within the 10-minute observation interval. A particularly serious drawback with time sampling occurs when a single instance of a behavior occurs

just before of just after the observer looks up to record the occurrence, resulting in a record of nonoccurrence.

When time sampling intervals are divided into segments by minutes as opposed to seconds, the procedure allows for longer periods between observations. It is therefore more practical for simultaneous teaching and data collection. Indeed, the interval may be set at 15, 30, 45 minutes, or more, allowing for observation throughout an entire day or class period. As the interval gets longer, however, the similarity between the data recorded and the actual occurrence of the behavior probably decreases (Saudargas & Zanolli, 1990). Time sampling is suitable primarily for recording behaviors that are frequent or of long duration, such as attention to task, out-of-seat behavior, or thumb sucking.

The longer the time sampling interval, the less accurate the data will be.

Although time sampling is practical for classroom use, its usefulness may diminish as a behavior-change program progresses successfully. For example, if Barry's teacher decides to use time sampling to record his out-of-seat behavior, she may observe that during baseline almost all of her observations indicate that Barry is out of his seat when she records at 15-minute intervals during the first 90 minutes of school. However, when a contingency is applied—"Barry, you will receive one token for each 15 minutes that you

FIGURE 3–12 ***Time sampling data sheet for a 3-hour observation period, with 15-minute intervals.***

Student:_____

Behavior:_____

_____

Date:_____

Time start:_____
Time end:_____
Total time period:_____

Observer:_____

CODE: X = occurrence O = nonoccurrence (note at the end of each interval)

15	30	45	00
15	30	45	00
15	30	45	00

Data Summary

Number of intervals of occurrence:
Percent of intervals of occurrence:
Number of intervals of nonoccurrence:
Percent of intervals of nonoccurrence:

stay in your seat"—the procedure will become less useful; for example, if the teacher holds to her recording procedure, records out-of-seat behavior only at the end of the interval, and provides the reinforcer if Barry beats the clock to his seat. If the procedure is successful in spite of this assumption, the data may reflect a complete absence of the behavior long before it has really been eliminated. Barry may be out of his seat for a short while a number of times that do not happen to coincide with the end of intervals. At this point, because the behavior is occurring less frequently and for shorter durations, event or duration recording may be as practical and much more accurate.

Data collected using interval recording or time sampling can be used to measure behavior along the frequency dimension by reporting the number of intervals during which the behavior occurred. These data, however, cannot be converted to rate. One cannot say that a certain behavior occurred at the rate of two per minute when what was recorded was the behavior occurring during two 10-second intervals in a 60-second period. Interval and time sampling data are most often expressed in terms of the percentage of intervals during which the behavior occurred. The procedure for converting raw data into percentages will be discussed in the next chapter. Measurement of duration can be approximated using interval recording, but this procedure does not lend itself to measures of latency. Force, locus, and topography may be measured, as with event recording, if they are included in the operational definition.

To recapitulate some of the important points and differences concerning interval recording and time sampling:

1. Both interval recording and time sampling provide an approximation of how often a behavior occurred. Neither is as accurate as event recording, which provides an exact count of occurrences.
2. Because interval recording divides an observation period into smaller intervals (usually seconds in length rather than minutes) than does time sampling, it provides the closer approximation to actual occurrence.
3. Interval recording is usually used for a short observation period (perhaps 15 minutes), whereas time sampling is used for longer periods of time (an entire morning).
4. Because time sampling divides an observation period into longer intervals, it is easier to manage while teaching.
5. The occurrence of a behavior is noted and recorded at any time during an interval when using interval recording. The occurrence of a behavior is noted and recorded only at the end of an interval when using time sampling.
6. For both interval recording and time sampling, the observer reports the number of intervals in which a behavior occurred (or did not occur), not the number of times it occurred. That information is not recoverable from data collected using these methods.

Ms. Simmons Observes Pencil Tapping

Ms. Simmons is an elementary-school consultant teacher for students with learning disabilities. One of her students, Arnold, tapped his pencil on the desk as he worked. He completed a surprising number of academic tasks but tapped whenever he was not actually writing. This behavior was very annoying to his general education teacher. Ms. Simmons had tried counting pencil taps but found that Arnold tapped so rapidly that the pencil was a blur.

She was working with another student, Shane, on paying attention and concentrating. She decided that recording interval data would be an excellent task for Shane. She carefully defined the behavior, provided Shane with a recording

sheet on a clipboard, gave him a stopwatch, and told him to mark a + in the box if Arnold tapped during a 10-second interval and a − if he didn't. She left Shane to observe and went to teach a small group. Very soon, she heard the sound of the clipboard hitting the floor and an expression of annoyance totally unacceptable in the classroom.

"I'm sorry, Ms. Simmons," said Shane, "but how the [explective deleted] am I 'posed to watch the kid, the watch, and the sheet at the same time?"

Realizing that she had asked too much of a student who had difficulty concentrating, Ms. Simmons revised her procedure. She took her cassette recorder home and prepared to make a tape that would "bleep" every 10 seconds. She set up the recorder next to her electronically controlled microwave oven and touched the "clear" button every 10 seconds as she watched the stopwatch. She had to start over only once when her teenage son, who had learned to expect dinner after a small number of "bleeps" (see Chapter 9 for a discussion of stimulus control), came in to ask when and what he was going to be fed.

The next day Ms. Simmons provided Shane with earphones and a 15-minute tape divided into 10-second intervals. Shane was able to record pencil tapping this way very efficiently, even though he couldn't hear it.

Variations on the Data Collection Sheet

Behavioral dimensions appropriate to interval and time sampling recording.

On the interval recording and time sampling data collection sheets presented so far, the teacher records the occurrence or nonoccurrence of a behavior for a single student. These basic data collection formats are, however, flexible tools that can be adapted easily to meet a variety of instructional situations. The most common adaptations are for: (1) a data sheet that is more descriptive of a target behavior whose operational definition may include several topographies (for example, stereotypic behavior defined as hand flapping, body rocking, and finger flicking); (2) a data sheet to accommodate taking data on more than one behavior at a time (for example, out of seat and talking); and (3) a data sheet to accommodate data for more than one student at a time.

Each of these data collection needs may be met by using coded data. Data collected simply to record the occurrence or nonoccurrence of a single behavior are called dichotomous data. The teacher is recording that the behavior either occurred or did not occur. For a fuller description of a behavior's various topographies or to record multiple behaviors, each behavior or variation is assigned a letter code to be used during data collection. If the observer wants to record data for multiple students, each student is assigned a letter code to be used on the data sheet.

There are at least three basic formats for coded data sheets: a legend coded data sheet, a prepared code data sheet, and a track coded data sheet. Each format can be used with either interval recording or time sampling.

Coded date sheet with a simple legend. A coded data collection sheet may have a legend like the one on a road map listing the behaviors being observed and a code for each. The example at the top of Figure 3–13 shows a few rows of data from an observation period during which interval recording was used to record occurrences of "disturbing one's neighbor." The legend includes codes for hitting (H), talking (T), and pinching (P). In this example, during the first minute of observation, Hector hit his neighbor during the first interval, talked during the third and fourth intervals, and talked and hit his neighbor during the fifth interval.

If the teacher wants to collect data on the same behavior being performed by more than one student, she may use this same format by providing a code for each student. For example, on the bottom of Figure 3–13 the teacher has made an adapted interval recording data sheet on which he will record occurrences of talking during the change of class

FIGURE 3–13 ***Legend coded data sheets for an interval recording observation.***

Student: Hector

Time Start: 9:10

Date: 9-18

Time End: 9:15

Observer: Ms. Hughes

Behavior Codes

H = hitting T = talking P = pinching

Minutes	10"	20"	30"	40"	50"	60"
1	H	—	T	T	T H	—
2	—	P	P	—	—	—
3	—	—	—	T	T	—
4	H	—	—	—	—	—
5	—	H	—	T H	T H	—

Student Jan, Ruth, Veena

Behavior Talking during change of class

Date 9-14

Time Start 11:05

Observer Mr. Nelson

Time End 11:10

Student Codes

J = Jan R = Ruth V = Veena

Minutes	10"	20"	30"	40"	50"	60"
1	JR	R	—	R	—	—
2	—	RV	—	RV		R
3	JR	—	RV	RV	—	—
4		—	—		—	—
5	—	—	—	RV	R	—

by Jan (J), Ruth (R), and Veena (V). In this example, during the first minute of observation, Jan and Ruth were talking during the first interval and Ruth was talking during the second and fourth intervals. Veena was not recorded as having talked until the second interval of the second minute.

Coded data sheet with a prepared format. A second type of coding, represented in Figure 3–14, uses letter codes for multiple behaviors or multiple students that are pre-entered in each cell (the space representing an interval) of the data sheet (Alberto, Sharpton, & Goldstein, 1979). With this prepared format the observer simply places a slash through the appropriate letter(s) indicating the behavior(s) that occurred or the student(s) who engaged in the behavior. The example at the top of Figure 3–14 shows rows of data from a 3-hour observation period during which the teacher used time sampling to record Sylvia's social interaction. The operational definition of social interaction included six potential elements, each with a code: I: student initiated the interaction, R: student responded to an initiation by someone else, S: the interaction was with another student, A:

FIGURE 3–14 ***Prepared coded data sheets for a time sampling observation.***

Student: Sylvia

Date: 11-6

Observer: Ms. Fannin

Behavior: social interaction

Time Start: 8:15

Time End: 11:15

Hour	10'			20'			30'			40'			50'			60'		
1	I	S	V	I	S	V	I	S	~~V~~	I	S	V	I	S	V	I	S	~~V~~
	R	A	P	R	A	P	~~R~~	~~A~~	P	R	A	P	R	A	P	~~R~~	~~A~~	P
2	I	S	~~V~~	~~I~~	S	V	I	S	V	~~I~~	S	V	~~I~~	S	V	I	S	V
	~~R~~	~~A~~	P	R	~~A~~	~~P~~	R	A	P	R	~~A~~	~~P~~	R	~~A~~	~~P~~	R	A	P
3	I	S	V	I	S	~~V~~	I	S	V	I	S	V	I	~~S~~	V	I	S	V
	R	A	P	~~R~~	~~A~~	~~P~~	R	A	P	R	A	P	~~R~~	A	~~P~~	~~R~~	~~A~~	~~P~~
4	I	S	V	I	S	V	I	S	V	I	S	V	I	S	V	I	S	V
	R	A	P	R	A	P	R	A	P	R	A	P	R	A	P	R	A	P
5	I	S	V	I	S	V	I	S	V	I	S	V	I	S	V	I	S	V
	R	A	P	R	A	P	R	A	P	R	A	P	R	A	P	R	A	P

I = initiate S = with student V = verbal

R = respond A = with adult P = physical

Student: Atal, Carmen, Kyle, Hanne

Date: 3-21

Observer: Mr. Klein

Behavior: Active engagement with task

Time Start: 9:00

Time End: 12:00

Period	10'		20'		30'		40'	
1	A	~~C~~	A	~~C~~	A	~~C~~	A	~~C~~
	~~K~~	H	K	H	K	~~H~~	K	~~H~~
2	A	~~C~~	A	~~C~~	A	~~C~~	A	~~C~~
	~~K~~	H	~~K~~	H	~~K~~	~~H~~	~~K~~	H
3	A	~~C~~	A	~~C~~	A	~~C~~	A	~~C~~
	~~K~~	H	~~K~~	H	~~K~~	~~H~~	~~K~~	~~H~~

A = Atal K = Kyle C = Carmen H = Hanne

the interaction was with an adult, V: the interaction was verbal, and P: the interaction was physical. In this example, during the first hour, Sylvia's first recorded social interaction was verbal in response to an adult, occurring at the end of the third interval.

If the observer wants to record the occurrence of a single behavior across several students, each student's code is placed in one cell. At the bottom of Figure 3–14, there is an adapted time sampling data sheet on which "active engagement with their task" was recorded for four students across three morning class periods. In this example, during the first period, Carmen and Kyle were actively engaged at the end of the first interval, Carmen at the end of the second, and Carmen and Hanne at the end of the third and fourth.

These data collection procedures are very flexible. They can be used to record data on several students simultaneously.

Coded data sheet using tracks. A third format for coding more than one behavior or for coding more than one student is the use of a tracking format (Bijou et al., 1968). The top of Figure 3–15 shows on-task behavior being recorded, but the teacher also wanted to know the general nature of any off-task behavior. Therefore, in addition to providing space to note the occurrence or nonoccurrence of on-task behavior, she provided track rows to indicate the general nature of any off-task behavior that occurred. The teacher would simply put a check mark in the appropriate cell(s) to indicate which behavior(s) occurred. This time sample data sheet indicates that at the end of the first and second 5-minute intervals, Rose was engaged in a motor off-task behavior; at the end of the third, fourth, and fifth intervals, she was on task; she was verbally and motorically off task at the end of the sixth; and she was passively off task at the end of the last two intervals. Alternatively, if the teacher wants to record data on two or three students simultaneously, the data sheet can be adapted by providing a row for each student's data.

Use codes to record data on several behaviors simultaneously.

Collecting data on students as a group. One way of adapting interval recording or time sampling for use with a group is the *round-robin* format (Cooper, 1981; Lloyd, Bateman, Landrum, & Hallahan, 1989). With this format the observer obtains an estimate of the group's behavior by observing and recording the behavior of a single group member during each interval. When conducting a language lesson, for example, the teacher might choose to monitor the group's attending behavior. As presented in Figure 3–16, the language period is divided into equal 15-second intervals to accommodate a group of four students, with the name of each group member assigned to each interval.

FIGURE 3–15 ***Track coded data sheet for time sampling.***

Student: Rose Date: 2-6

Observer: Ms. Paster Time Start: 10:40

Behavior: on task

	5'	10'	15'	20'	25'	30'	35'	40'
on task			X	X	X			
verbal off task						X		
motor	X	X				X		
passive							X	X

FIGURE 3–16 ***Round-robin format of interval recording.***

	1st 15-second interval	2nd 15-second interval	3rd 15-second interval	4th 15-second interval
	Kate	Michael	Harry	Jody
1				
2				
3				
4				

In this example, Kate is observed for occurrence or nonoccurrence of attending during the first 15-second interval of each minute, Michael's attending is observed and recorded during the second interval of each minute, Harry's during the third, and Jody's during the fourth. Because each student is observed during only one of the intervals per minute, on a round-robin basis, the resulting data provide a representation of the whole group's attending behavior, but not an accurate representation of the attending behavior of any single member of the group. Another data collection method must be used to focus on individual students. The round-robin format may also be used when collecting data on a whole class. In order to measure on-task behavior, Sutherland, Wehby, and Copeland (2000) observed each of four rows of students in rotation. Using time sampling, the teacher noted at the end of each interval whether all the students in the selected row were oriented toward the appropriate task or person and thus, as operationally defined, on task. Using random order, the teacher observed each row of students several times during the observation period.

Duration and Latency Recording

Event recording, interval recording, and time sampling collection techniques focus primarily on exact or approximate counts of the occurrence of a behavior. Duration and latency recording differ from these systems in that the focus is on a temporal rather than a numerical dimension of the behavior.

Duration and latency recording emphasize measures of time rather than instances of behavior.

Duration Recording

Duration recording is used when the primary concern is the length of time a student engages in a particular behavior. For example, if a teacher wants to know about a student's out-of-seat behavior, either event recording or duration recording might be appropriate. Event recording would provide information about the number of times a student left her seat. If, however, the teacher's concern is *how long* she stays out of her seat, the most appropriate data collection method would be duration recording. In this example, event recording would mask the temporal nature of the target behavior. Although event data might indicate that the number of times the student left her seat had decreased substantially, it would not reveal that the length of time spent out of seat might actually have increased.

You could measure average duration of tantrums, of time spent on task, or of recreational reading.

Duration recording, like event recording, is suitable for behaviors that have an easily identifiable beginning and end. It is important to define clearly the onset of the behavior and its completion. Using clearly stated operational definitions, researchers have measured the duration of academic engagement/time on task (Kim & Hupp, 2007; Lane et al., 2003; Regan, Mastropieri, & Scruggs, 2005); in-seat behavior (Roane, Fisher, & McDonough, 2003); appropriate social interactions (Chin & Bernard-Opitz, 2000; Shukla, Kennedy, & Cushing, 1999); vocational skills such as sweeping floors, bussing tables, stuffing envelopes, and filing (Grossi & Heward, 1998; Worsdell, Iwata, & Wallace, 2002); leisure activity (Stewart & Bengier, 2001; Zhang, Gast, Horvat, & Dattilo, 2000); inappropriate behaviors such as finger or thumb sucking (Ellingson et al., 2000; Stricker, Miltenberger, Garlinghouse, Deaver, & Anderson, 2001); aggressive behaviors such as biting, kicking, scratching, pushing, hitting, and spitting (Oliver, Oxener, Hearn, & Hall, 2001; Romaniuk et al., 2002); and periods of disruptive behavior such as yelling, cursing, and taking the property of others (Gresham, Van, & Cook, 2006).

The observer may time the duration of the behavior using the second hand of a watch or wall clock, but a stopwatch makes the process much simpler. For certain behaviors such as seizures or tantrums, an observer may use an audio or video recorder. The duration of the episode may be determined later using the resulting permanent product and a stopwatch or the automatic timer on the VCR.

Total duration could be used to record the time spent talking, reading, or playing with toys.

There are two basic ways to record duration data: average duration and total duration. The average duration approach is used when the student performs the target behavior routinely or with some regularity. In a given day, the teacher measures the length of time consumed in each occurrence (its duration) and then finds the *average duration* for that day. If the behavior occurs at regular but widely spaced intervals (for example, only once per day or once per class period), the data may be averaged for the week. One behavior that can be measured by duration data is time spent in the bathroom. Perhaps his teacher feels that each time John goes to the bathroom, he stays for an unreasonable length of time. To gather data on this behavior, she decides to measure the amount of time he takes for each trip. On Monday, John went to the bathroom three times. The first trip took him 7 minutes, the second 11 minutes, and the third 9 minutes. If she continued to collect data in this manner during the rest of the week, the teacher would be able to calculate John's average duration of bathroom use for the week. *Total duration recording* measures how long a student engages in a behavior in a limited time period. This activity may or may not be continuous. As an example, the target behavior "appropriate play" might be observed over a 15-minute period. The observer would record the number of minutes the student was engaging in appropriate play during this period. The child might have been playing appropriately from 10:00–10:04 a.m. (4 minutes), from 10:07–10:08 a.m. (1 minute), and from 10:10–10:15 a.m. (5 minutes). Although such a behavior record is clearly noncontinuous, these notations would yield a total duration of 10 minutes of appropriate play during the 15-minute observation period.

Latency Recording

Latency recording measures how long a student takes to begin performing a behavior once its performance has been requested. This procedure measures the length of time between the presentation of an antecedent stimulus and the initiation of the behavior. For example, if a teacher says, "Michael, sit down" (antecedent stimulus) and Michael does, but so slowly that 5 minutes elapse before he is seated, the teacher would be concerned with the latency of the student's response. Latency recording has been used to measure

the time between an antecedent instruction and students' beginning to put away toys (Shriver & Allen, 1997), beginning academic assignments (Hutchinson, Murdock, Williamson, & Cronin, 2000; Maag & Anderson, 2006; Wehby & Hollahan, 2000), beginning the steps of self-catheterization (McComas et al., 1999), initiating transitions (Ardoin, Martens, & Wolfe, 1999), between incidents of pica (Pace & Toyer, 2000), and latency to the onset of destructive behavior as a means of escaping academic and self-care tasks (Zarcone, Crosland, Fisher, Worsdell, & Herman, 1999).

As seen in Figure 3–17 a basic collection sheet for duration or latency data should provide information on the temporal boundaries that define the procedures. A duration recording data collection sheet should note the time the student began the response and the time the response was completed. A latency recording data collection sheet should note the time the student was given the cue to begin a response (antecedent stimulus) and the time he actually began to respond.

Duration and latency recording are closely matched to the behavioral dimensions of duration and latency. Consideration of topography, locus, and force may also apply here. For example, a teacher might want to measure

- how long Calvin can perfectly maintain a position in gymnastics
- how long Rosa talks to each of a number of other students
- how long after being given a nonverbal signal to lower her voice Ellen actually does so
- how long David maintains sufficient pressure to activate a microswitch.

FIGURE 3–17 ***Basic formats for latency and duration recording data sheets.***

Latency Recording Data Sheet

Student: EDITH Observer MR HALL

Behavior: TIME ELAPSED BEFORE TAKING SEAT

Operationalization of behavior initiation: _____

Date	Time		Latency
	Delivery of S^D	Response initiation	

Duration Recording Data Sheet

Student: SAM Observer MS. JAMES

Behavior: TIME SPENT IN BATHROOM

Behavior initiation: _____

Behavior termination: _____

Date	Time		Duration
	Response initiation	Response termination	

How Can All This Be Done?

Thinking about all that goes on in a classroom can make the tasks in this chapter seem overwhelming. In fact, they may be headed for the pile of advice that is "just not practical" in my class of 34 students in freshman English with 6 "inclusion students" of varying disabilities, or of 14 students with behavioral disorders and hyperactivity, or of 6 students with severe mental retardation or autism. Before tossing this chapter aside, remember that we acknowledged in the first paragraph that some of this may not be practical on a daily basis. Knowing this content, however, will enable you to design an appropriate data-based accountability system and will enable you to read educational research with more authority and be better prepared to apply it in your classroom. Data collection should be a tool that contributes instructional logic to classroom management. Data collection provides the basis for selecting appropriate objectives, arranging instructional groupings, and meeting the requirements of accountability. Well, then, how *can* all this be done? Here are some suggestions about how much data to collect, who should collect it, and what kind of help technology can provide.

One question to ask about how much data to collect is: How much data should I collect within an instructional session? One can either collect trial-by-trial data (recording *all* occurrences of the behavior) or collect probe data (recording only a *sample* of the occurrences). Trial-by-trial data collection during *instruction* records whether every response during the session is correct or incorrect. Teachers can collect probe data (data on some, but not all, responses) in two ways. It may be collected just before or just after an instructional session. If the teacher is about to conduct a 20-minute lesson on multiplying by 6, she may probe or sample student knowledge before beginning the lesson (measuring what the student has maintained since the last lesson), or she may probe just after the lesson (measuring what the student just learned) by asking the student to multiply a sample of numbers by 6. The term *probe* is also used to mean recording correct or incorrect use of a target behavior in an untrained setting (multiplying by 6 in the grocery store when shopping for a group) or within an untrained variation of the response (in a word problem).

During *behavior-change programs*, trial-by-trial data collection means continuous data recording for the entire period the contingency is in place—for example, recording instances or intervals of verbal aggression during all of PE or recording intervals of attending during the whole general education science class. There are several ways to use probe data with behavior change programs: (a) a contingency may be in effect all day, but student behavior is sampled by collecting data only during specific time periods—the first 10 minutes of each class period, the first five times the student has the opportunity to greet someone verbally, or (b) students' behavior may be probed or sampled during time periods when the contingency was not initially taught, or (c) students' behavior may be probed or sampled in environments where the contingency was not initially taught.

Another question to ask is: How often should I collect data during a week? Some advice is provided in the professional literature. Data collected once every 2 or 3 days closely approximates data collected daily (Bijou, Peterson, Harris, Allen, & Johnston, 1969). Fuchs and Fuchs (1986) reviewed studies in general education and special education in which formative evaluation was used for academic learning with preschool, elementary, and secondary students. They concluded that systematic formative evaluation was effective whether data were recorded daily or twice per week. Farlow and Snell (1994) suggested when implementing a new behavior program or teaching a new skill, data should be collected every day or every teaching session until the student shows steady progress over six data points or 2 weeks. At that point accurate and reliable

judgments of ongoing progress can be made with data collected twice a week. When a learning problem is suspected, data should be recorded daily while programmatic adjustments are made. Once progress is seen over approximately 2 weeks, then data collection twice a week is sufficient.

The teacher is not the only one who can or should collect data. You can get someone to help collect data. Counting by ones is not a higher-order skill. Appropriate training and practice, however, is essential to reliable data collection.

In a cotaught class, the special and general educator can collaborate on data collection to monitor progress of all the students in the class. In classrooms and community settings the teacher and paraeducators can all collect data. A speech–language pathologist can collect data on communication objectives, and a physical or occupational therapist can collect data on a student's learning to self-catheterize or climb stairs. In special and general education classes, especially classes in which students with disabilities are included, peers can be trained as data collectors in dyads and small groups (Marchand-Martella, Martella, Bettis, & Blakely, 2004; Simmons, Fuchs, Fuchs, Hodges, & Mathes, 1994). Students should record data on themselves whenever possible. The ability to self-assess through recording one's own behavior is a component of independence. This is explored in detail in Chapter 11.

There is low-tech and high-tech assistance available for data collection. We have already described a variety of low-tech options that range from paper clips and golf counters for event recording, kitchen timers and stopwatches for duration recording, audiotapes with prerecorded chimes, and watches that chime and vibrate at specific intervals. Whenever possible we suggest making a permanent product with videotape. It is always easier and more accurate to record data on a high-frequency or disruptive behavior later while watching it on videotape rather than when trying to manage the behavior at the same time.

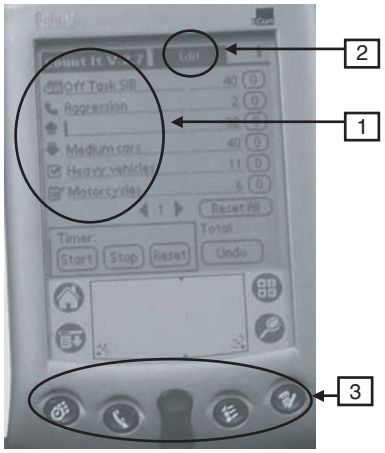
High-tech computerized systems have greatly advanced the ease and accuracy of data collection. Kahng and Iwata (1998) described and reviewed key characteristics of 15 computerized data collection systems that use primarily laptop computers. They suggested these systems can improve reliability and accuracy of behavioral recording compared to paper-and-pencil methods. Most of the systems reviewed use IBM-compatible software; five use MacOS. The systems have a range of capabilities from those that can collect frequency, interval, time sampling, duration, and latency data (e.g., Behavioral Evaluation Strategy & Taxonomy, The Behavior Observer System, or The Direct Observation Data System) to those with a limited range, such as the Ecobehavioral Assessment System Software, which can collect interval data, or the Social Interaction Continuous Observation Program for Experimental Studies, which can collect frequency and duration data. Most of the systems include data analysis programs, and about a third include a program to compute interobserver agreement.

A potential concern when using a computerized recording system in the classroom and other applied settings is the introduction of obtrusive materials that are likely to add confounding variables to a natural environment. Such materials increase the likelihood of students' reactivity to novel materials and may change students' behavior. Research procedures that are obtrusive produce rival explanations and conclusions (Krathwohl, 1998).

A possible method of collecting observational data while remaining unobtrusive is to use a personal digital assistant (PDA), commonly referred to as a handheld computer, such as the "Palm." PDAs provide the teacher and researcher with portability, ease and versatility of data recording, and immediate data analysis. The data can be exported form the PDA to a desktop computer. With software such as Microsoft Excel or Lotus 123, the data can be converted into spreadsheets, graphs, or tables for further analysis.

An example of a software program that can be used in the classroom is *Count It* (Molgaard, 2001), a shareware database that can be downloaded easily to a PDA. *Count*

FIGURE 3–18 ***Using the Count It shareware to collect event data.***



1 = counters, 2 = edit, and 3 = hardware

Instructions for Event Recording

There is an option of recording one target behavior or dependent variable at a time, or multiple dependent variables. First, label *Counters* with a specific target behavior by selecting *Edit*. Next, delete the existing labels, type in the new labels, select *Done*, and then *Edit*. The hardware buttons on the PDA correspond to the specified label. Repeat the step until all target responses or dependent variables are cataloged. During the observation, press the hardware button that denotes the specific response. The counter will automatically increase the dependent variable by one. At the end of the observation the total frequency for each response is displayed.

Adapted by Cihak & Alaimo, 2003. Used by permission.

It can file individual behaviors and offers functions for customized data recording. The online manual provides detailed explanations, and the author is accessible through electronic mail. Figure 3–18 provides instructions for the use of *Count It* for event recording that were developed for use in general and special education classes in a local school system (Cihak & Alaimo, 2003).

Summary of Data Collection Systems

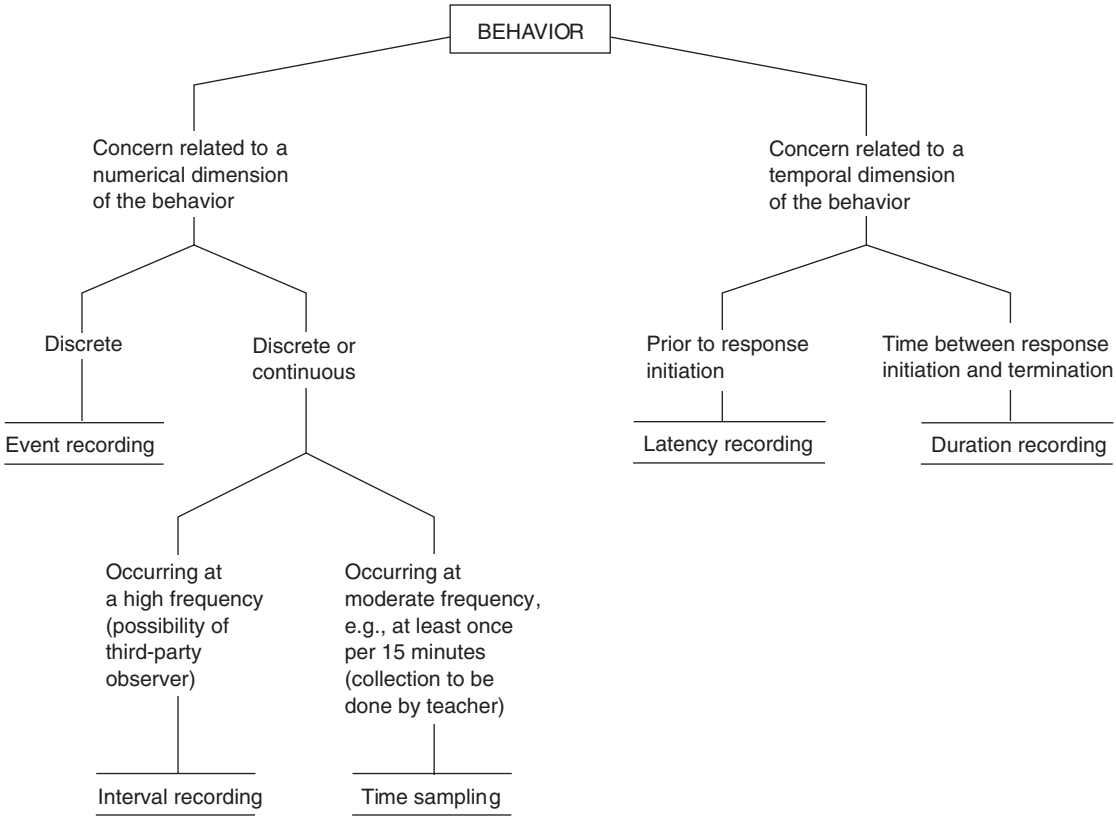
Questions for the data collector.

How to choose a recording system.

The five observational systems available to the data collector are event recording, interval recording, time sampling, duration recording, and latency recording. Figure 3–19 summarizes the decision making involved in selecting the system appropriate for a particular target behavior. This process is based on a series of questions to be answered by the data collector:

1. Is the target behavior numerical or temporal?
2. If it is numerical:
   a. Is the behavior discrete or continuous?
   b. Is the behavior expected to occur at a high, moderate, or low frequency?
   c. Will I be able to collect data during intervention or instruction, or will I need a third party to collect the data so as not to interrupt instruction?
3. If it is temporal, do I want to measure the time before initiation of the response or time elapsed during performance of the response?

FIGURE 3–19 ***Selected observational recording procedures.***



RELIABILITY

Computing reliability for permanent product data.

When data collection depends on human beings, there is always the possibility of error. Even in the case of permanent product data, which are easiest to record, mistakes may happen. Teachers occasionally count math problems as incorrect even when they are correctly done or overlook a misspelled word in a paragraph. Because there is something tangible, however, the teacher can easily recheck the accuracy or **reliability** of her observations of the behavior. In using an observational recording system, however, the teacher does not have this advantage. The behavior occurs and then disappears, so she cannot go back and check her accuracy. To be sure the data are correct, or reliable, it is wise to have a second observer simultaneously and independently record the same behavior periodically. When this is done, the two observations can be compared and a coefficient or percent of *interobserver reliability*, or *interobserver agreement*, may be computed (Baer, 1977; Johnston & Pennypacker, 1993; Sidman, 1960).

Computing reliability for event recording.

To check event recording, the teacher and a second observer, a paraprofessional or another student, simultaneously watch the student and record each instance of the target behavior. After the observation period, the teacher calculates the coefficient of agreement or reliability by dividing the smaller number of recorded instances by the larger number of recorded instances. For example, if the teacher observed 20 instances of talking out in a 40-minute session, and the second observer recorded only 19, the calculation would be 19/20 = 0.95. Therefore, the coefficient of interobserver agreement would be 0.95, or 95%. This method of calculating reliability for research purposes lacks a certain amount

of precision and has therefore been referred to as a gross method of calculation. "The problem is that this method does not permit the researcher to state that both observers saw the same thing or that the events they agreed on were all the same events" (Tawney & Gast, 1984, p. 138). In other words, there is no absolute certainty that the 19 occurrences noted by the paraprofessional were the same ones noted by the teacher.

The reliability of duration and latency data is determined by a procedure similar to that of event recording, except that the longer time is divided into the shorter, as in the following equation:

$$\frac{\text{Shorter number of minutes}}{\text{Long number of minutes}} \times 100 = \text{percent of agreement}$$

When using interval recording or time sampling, the basic formula for calculating reliability is:

$$\frac{\text{Agreements}}{\text{Agreements + Disagreements}} \times 100 = \text{percent of agreement}$$

Computing reliability for duration and latency recording.

If the data shown represent 10 intervals during which the teacher and the paraprofessional were recording whether or not Lauren was talking to her neighbor, we see that their data agree in 7 intervals (that is, intervals 1, 2, 3, 4, 6, 7, and 8); their data are not in agreement in 3 intervals (that is 5, 9, and 10). Therefore, using the basic formula, the calculation for reliability would be as follows:

$$\frac{7}{7 + 3} \times 100 = 70\%$$

	1	2	3	4	5	6	7	8	9	10
Teacher	X	X	-	-	X	X	-	-	X	-
Paraprofessional	X	X	-	-	-	X	-	-	-	X

Under certain research circumstances an additional, more rigorous determination of reliability should be considered. This should be a calculation of *occurrence reliability* or *nonoccurrence reliability.* When the target behavior is recorded to have occurred in less than 75% of the intervals, occurrence reliability should be calculated. When the target behavior is recorded to have occurred in more than 75% of the intervals, nonoccurrence agreement should be computed (Tawney & Gast, 1984). These coefficients are determined with the same basic formula (agreements/[agreements + disagreements] × 100), except only those intervals in which the behavior occurred (or did not occur) are used in the computation.

Factors That May Affect Data Collection and Interobserver Agreement

In general, applied behavior analysts aim for a reliability coefficient of around 0.90. Anything less than 0.80 is a signal that something is seriously wrong. A low coefficient of reliability can often be explained by examining the operational definition of the behavior; sloppy definitions, those that do not expressly state a behavior's topography or when it begins and ends, result in low reliability. The observers may not have been told exactly what they are to observe. Insufficient agreement may also be due to a lack of sufficient training in the data collection system. Either the primary or secondary

observer may not be employing the mechanics of the data collection system correctly, resulting in differing records of the occurrence of the behavior. The environment in which data is collected may also be a factor. In natural settings such as classrooms, homes, communities, and workplaces, many variables may affect behavior and many behaviors may be occurring simultaneously. Given all that may be going on in natural settings, an observer who is unfamiliar with the setting and collects data only occasionally for reliability purposes will be less at ease and possibly more distracted during data collection, and her data may be less accurate than that of an observer who is regularly in that setting (Fradenburg, Harrison, & Baer, 1995; Repp, Nieminen, Olinger, & Brusca, 1988).

Kazdin (1977) suggested four sources of bias that can also affect interobserver agreement: reactivity, observer drift, complexity, and expectancy.

Reactivity: As teachers are well aware, the presence of an observer can affect the behavior of both the students being observed and of the teacher. This effect is known as *reactivity* (Repp et al., 1988). A student who knows that he is being observed may react by being very "good" or may put on an unruly show, either of which will give a false view of a target behavior. Some teachers give more prompts to a target student when an observer is present (Hay, Nelson, & Hay, 1977), and some increase their rate of instruction and positive feedback, both of which may affect the typical occurrence of the behavior (Hay, Nelson, & Hay, 1980). Simply knowing another person is present collecting reliability data can influence the accuracy of the primary observer. Such knowledge has influenced reliability data by as much as 20 to 25%. It is suggested that reliability checks should be unobtrusive or covert, if possible, or that the second observer collect data on several students including the target student, or that the second observer be someone familiar to the student, such as a classroom paraprofessional. Some of these suggestions may not be practical in every instance, but just limiting communication between the first and second observer during the observation period can reduce their influence on one another's observations.

Observer Drift: Observer drift is the tendency of observers to change the stringency with which they apply operational definitions. Over time an observer may reoperationalize the definition as it becomes less fresh in his mind. The observer may begin to record as "instances" behaviors that do not exactly conform to the operational definition. If the operational definition appears on every copy of the data sheet, the observer can easily consult it. Observers should periodically review definitions together and conduct practice sessions during the course of the program.

Complexity: A third influence on the reliability of data concerns the complexity of the observational coding system. The more complex the system, the more the reliability is in jeopardy. Complexity refers to the number of different types of a response category being recorded (for example, the number of types of disruptive behavior being observed simultaneously), the number of different students being observed, or the number of different behaviors being scored on a given occasion. In a classroom, the teacher may mitigate the effects of complexity by limiting the number of behaviors or students observed at any given time.

Kazdin (1977) provides suggestions for limiting complexity bias in research studies.

Expectancy: The fourth bias is that of expectancy. Observers' preconceived notions about students based on their past experiences with them or on information from parents or previous teachers have the potential to bias their interpretation of what they are seeing. In addition, when observers are teachers who expect behavior change (because of the terrific job they did with their intervention), they are likely to find it. The reverse is also true; a teacher who has decided that nothing can be done with a student is not likely to see a change in behavior accurately.

Observers may be biased because of the student's sex, race, appearance, or previous history. In addition, a bias may result from the purpose of the observation (Repp et al., 1988). Teachers may be biased data collectors if the failure of a behavior-change strategy will result in a problem student's being moved to a different environment.

The procedures described in this section are adequate to determine reliability for most teachers, especially if efforts are made to control bias. More stringent standards are sometimes applied in research studies. The student who is interested in learning more about interobserver reliability should consult Hawkins and Dotson (1975) and the series of invited articles on this topic in Volume 10 (1977) of the *Journal of Applied Behavior Analysis*.

SUMMARY

We described the various dimensions of behavior (i.e., frequency, rate, duration, latency, topography, focus, and locus) and their relationship to data collection. Data collection procedures discussed included anecdotal reports, permanent product recording, and various observational recording systems (event recording, interval recording, time sampling, duration recording, and latency recording). To assist teachers to increase the accuracy of their data, procedures for determining interobserver agreement were outlined. As professionals responsible for student learning, teachers collect data to make determinations about the success of or need for change in their instruction or behavioral intervention.

KEY TERMS

permanent product recording
observational recording systems
event recording
interval recording
time sampling
duration recording
latency recording
discrete behaviors
frequency
controlled presentations
trial
reliability

DISCUSSION QUESTIONS

1. Jerry's behavior in his fifth-grade class was reported as "disruptive." The consulting teacher visited his classroom to collect some initial referral data. (a) She went into his class for 30 minutes on 3 days to count instances of "disruptive" behavior. (b) On 3 days she checked every 20 minutes between 9 a.m. and noon to see if he was being disruptive. (c) For 1 hour on Tuesday morning and a Thursday afternoon, she sat in Jerry's class and wrote down everything he did, what his teacher did, and significant actions of other students. What observational recording system did she use in each instance?
2. Susan never gets her math problems done before the end of class. To help determine the nature of her problems, the teacher could (a) give her a set of problems and record how long it was before she began to work, or (b) record how long it took her to complete the set of problems once she had begun. What recording system is being used in each instance?
3. Four student data collectors were observing John, a fourth-grade student. John was doing poorly in spelling. Observer 1 divided his observational time into 15-second intervals and noted whether John was working in his spelling workbook during each interval. Observer 2 went to John's desk at the end of the spelling period and counted the number of answers John had written in his spelling workbook. Observer 3 counted each time John put his pencil on the workbook and wrote something. Observer 4 divided the period into 5-minute

intervals and recorded whether John was working in his spelling workbook at the end of each interval. What recording procedure is each observer employing?

4. Mrs. Carrington wanted the students to help her check their knowledge of multiplication facts. The students were divided into pairs in order to ask each other the 7, 8, and 9 multiplication table facts and record their accuracy. Each student was given a packet of flash cards that had the problem statement on one side and the answer on the back of the card. Also, on the back was a place to mark whether the answer given by his or her partner was correct or incorrect. What recording procedure is being used by the students ?

5. Describe the various dimensions of the following behaviors.

   a. mutual toy play
   b. writing in a daily journal
   c. kicking furniture
   d. cleaning the glass doors in the frozen-food section of the grocery
   e. writing the letters of the alphabet
   f. riding a tricycle
   g. using a mouse to select the correct answer on a computer screen
   h. completing a sheet of long-division problems
   i. inititiating social greetings
   j. flicking fingers in front of one's eyes

6. The following is an anecdotal report of one session of community-based vocational instruction. Todd, his classmate Lucy, and their teacher were at Pets-Are-Us. The session's task was to move 4-pound bags of birdseed from the storeroom to the shelves at the front of the store. Transpose the information in the anecdotal report into the A-B-C column format.

   May 3, 9:20 a.m.: Teacher, Todd, and Lucy are in the storeroom. Teacher explains the task. She tells both students to pick up a bag and follow her. They do, and place bag on proper shelf. She leads them back to storeroom. Teacher tells Todd to pick up a bag of seed; he walks away. She tells him a second time. Teacher picks up a bag and takes Todd by the hand and walks out to the shelf. She hands him the bag and points to where it belongs; he puts the bag on the shelf. She tells him to go back to storeroom for another. In the storeroom she tells him to pick up a bag from the pile. The third time he is told, he picks one up and goes out front and puts a bag on shelf. On the way out to the floor with the next bag, Todd stops at a birdcage, drops bag, and begins to talk to birds. Several minutes later teacher comes for him. He ignores her. She puts his hands on bag then leads him to shelf. She then takes him to storeroom. He refuses to lift a bag. She hands him one. He drops it on floor. This is repeated twice. She takes a bag and leads him back out to the shelf. She tells him to go back to the storeroom. She goes to check on Lucy. Ten minutes later she finds Todd sitting on floor eating candy from his fanny pack. She takes the candy, tells him it is for later. She tells him again to go to the storeroom. When she looks for him again, he is at the rabbit cage. She leads him back to storeroom. Tells him to pick up bag. After third delivery of instruction teacher holds bag in front of him; he doesn't move his arms. She places his arms around the bag. He lets it drop through his hands and it splits open. She scolds. She goes to get broom. She returns, and he is sitting on the floor eating the birdseed. Teacher tells Todd, "Your behavior is not acceptable. Therefore, you will no longer be allowed to work today. Sit over there and time yourself out until we leave. I am very disappointed in your work behavior today."

Chapter 4

Graphing Data

Did you know that . . .

- A picture is worth a thousand data points?
- Connecting the dots is not just child's play?
- Graphs can be used as a communication tool?
- The same set of data can be graphed in more than one way?
- Graphing hasn't changed much since you learned to do it in elementary school?

Chapter Outline

Data collection, as you can imagine, results in a pile of data sheets. For data to be useful, the contents of those sheets must be rearranged in a way that allows them to be easily read and interpreted. The most common method of arranging and presenting data is to use a graph. A properly drawn graph provides a picture of progress across the time of instruction or intervention. Graphs should be simple and uncluttered but provide sufficient information to monitor progress.

For step-by-step instructions for using the computer to create graphs, refer *Graphing in Excel for the Classroom: A Step-by-Step Approach* (Cihak, Alberto, Troutman, & Flores, 2005).

Graphs serve at least three purposes. First, they provide a means for organizing data during the data collection process. Tallies on sheets of paper or coded entries on data collection forms are difficult, if not impossible, to interpret. Translating raw data into a graph provides an ongoing picture of progress (or lack thereof) that is easier to understand than thumbing through piles of data sheets. Second, an ongoing picture makes possible formative evaluation, the ongoing analysis of the effectiveness of an intervention. Formative evaluation makes it possible to see how well a procedure is working and to make adjustments if it is not working well. When the intervention is finished, inspecting a graph allows for summative program evaluation, the end result of an intervention or series of interventions. Third, graphs serve as a vehicle for communication among teacher, student, parents, and related service professionals. A properly constructed graph shows all the information about how the target behavior changes during an intervention. One should be able to read and understand the graph without having to read a prose explanation. The information shown on graphs can be used to write and evaluate progress reports, individualized education programs, and behavior management plans.

Purposes for graphs.

THE SIMPLE LINE GRAPH

BASIC ELEMENTS OF THE LINE GRAPH

Line graphs are commonly used to display data in a serial manner across the duration of instruction or intervention. This allows for ongoing monitoring of the behavior and evaluation of the instruction or intervention. Graphs can be constructed using graph paper or a computer program. The grid on the graph paper or the computer software makes it possible to plot data accurately, ensuring proper alignment and equal intervals among data points. When data are presented formally, as in publications, the grid is usually omitted. The following is a description of basic elements and conventions for constructing a simple line graph (*Journal of Applied Behavior Analysis,* 2000; Poling, Methot, & LeSage, 1994; Tawney & Gast, 1984). These are illustrated by the two graphs in Figure 4–1.

Axes

These axes don't chop wood.

A graph is constructed within a set of boundaries. These boundaries are called *axes* (*axis*, singular). A line graph has two axes: the horizontal **abscissa,** or *x axis*, and the vertical **ordinate,** or *y axis*. When the graph is completed, these axes are drawn in a ratio of 2:3. Thus, if the *y* axis is 2 inches long, the *x* axis should be 3 inches. If the *y* axis is 4 inches, the *x* axis should be 6 inches.

1. *Abscissa:* The abscissa is the horizontal line that serves as the bottom boundary of the graph. It shows how frequently data were collected during the period represented on the graph. It may be labeled as, for example, days, dates, or sessions. If sessions are used, it is helpful to provide some definition of the session, for example,

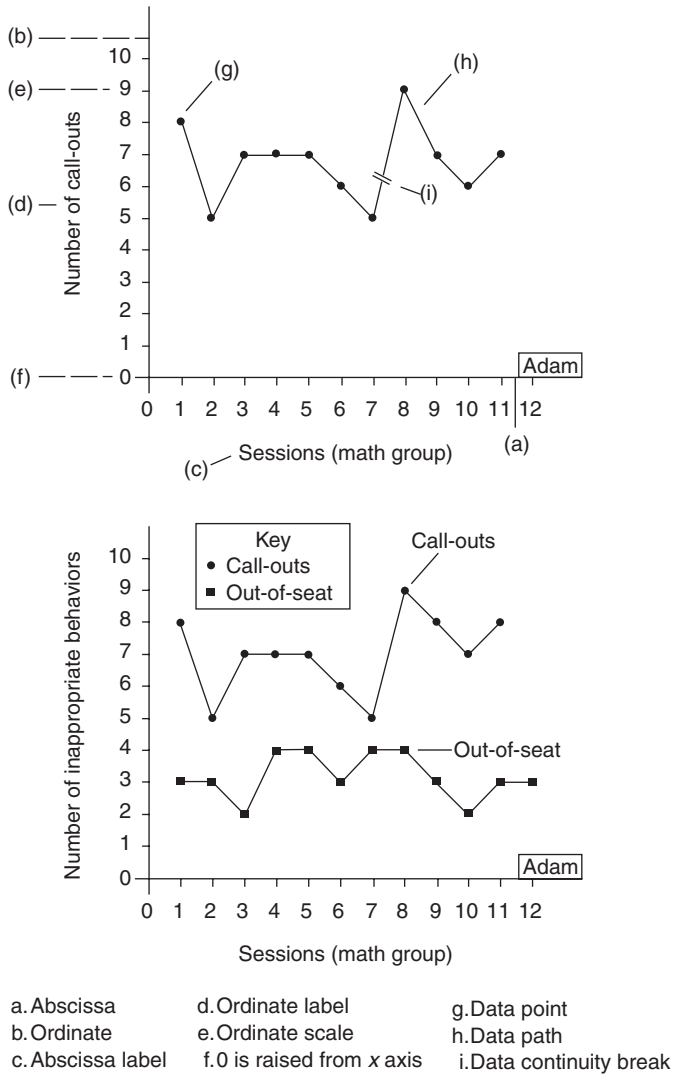


FIGURE 4–1
Basic elements and conventions for time graphs.

"sessions (9–9:40 a.m.)" or "sessions (math group)." The right boundary of the graph ends at the last session number.

2. *Ordinate:* The ordinate is the vertical line that serves as the left-hand boundary of the graph. The label on the ordinate identifies the target behavior and the kind of data that is being reported. For example, a label might read "number of occurrences of cursing," "rate of cursing," "number of intervals of cursing," or "percent of intervals of cursing." Standard data conversion procedures are presented in Table 4–1.

   a. *Ordinate scale:* The scale on the *y* axis, used to record the performance of the target behavior, always begins at 0. If one is reporting the number of occurrences of the behavior or number of intervals during which the behavior occurred, the scale begins at 0 and goes as high as needed to accommodate the largest number. This number is sometimes difficult to predict, and the

An advantage to computing percentage.

TABLE 4–1
Summary of data conversion procedures.

Type of Recording	Data Conversion	
Permanent product recording	Report number of occurrences . . .	if both time and opportunities to respond are constant.
Event recording	Report percentage . . .	if time is constant (or not of concern) and opportunities vary.
	Report rate . . .	if both time (which is of concern) and opportunities vary, OR if time varies and opportunities are constant.
Interval recording	Report number of intervals . . .	if constant.
Time sampling	Report percentage of intervals . . .	during or at the end of which behavior occurred.
Duration	Report number of seconds/ minutes/hours . . .	for which the behavior occurred.
Latency	Report number of seconds/ minutes/hours . . .	between antecedent stimulus and onset of behavior.

x axis = abscissa
y axis = ordinate.

researcher may have to redraw the graph if data are being plotted before the intervention is completed. If percent is being reported, the scale always goes from 0 to 100%. The scale may progress by single digits or by twos, fives, tens, or other multiples in order to accommodate the data. It makes the graph easier to read if the beginning point of the scale (the zero value) is raised slightly from the abscissa; data points are more easily discerned when they do not rest on the *x* axis.

b. *Scale break:* Occasionally, the ordinate scale may not be continuous. For example, if all the data points are above 40%, the bottom part of the graph will be empty and the top will be unnecessarily crowded. It is permissible to begin the scale at 0, draw two horizontal parallel lines between the first and second lines on the graph paper, and label the second line 40%.

Data

1. *Data point:* Small geometric forms, such as circles, squares, or triangles, are used to represent the occurrence of the target behavior during a time segment. For example, in the first graph shown in Figure 4–1, the student engaged in cursing eight times during session one; therefore, a data point is placed at the intersection of 8 on the *y* axis and 1 on the *x* axis. Each data point is independently plotted on the graph. The placement or value of one data point does not affect the placement or value of the next data point.

2. *Data path:* When a solid line is drawn connecting the data points, it forms the data path.
   a. A single geometric form is used to represent each point on a single data path.
   b. When more than one set of data appears on a graph, each is represented by a different geometric symbol. Which behavior is represented by each symbol and data path may be shown in one of two ways, both of which can be observed on the second graph of Figure 4–1. Each path may be labeled and an arrow drawn from the label to the path, or a legend (or key) may be provided listing each geometric symbol and the behavior it represents. No more than three different data paths should be plotted on a single graph. Additional graphs should be used when more than three data paths are necessary.
   c. The solid line of a data path implies continuity in the data collection process. If there is a break in the expected sequence of intervention (a student is absent; a special event occurs) and a regularly scheduled session does not occur, two parallel hash marks are placed on the data path to indicate the continuity break.
3. *Student (participant) identification:* The name of the student(s) is placed in a box in the bottom right corner of the graph.

Graphs with more than three data paths look cluttered.

Transferring Data to a Graph

Transferring Permanent Product Data to a Graph

Graphing permanent product data.

Permanent product data are reported as a number of items or a percentage of items resulting from behavior. For example, a teacher might record the number of math problems completed, the percent of correctly spelled words, the number of cans placed on a display shelf, or the number or dirty clothes placed in the hamper. If the number of opportunities for responding remains constant—as in spelling tests that always have 20 items or in a series of math worksheets that always have 10 problems—the data may be graphed simply as the number of items. If the number of opportunities varies—different numbers of test items or math problems—the teacher must calculate percentages (see Figure 4–2).

We calculate the percentage of correct responses by dividing the number of correct responses by the total number of responses and multiplying the result by 100, as shown:

$$\frac{\text{Number of correct responses}}{\text{Total number of responses}} \times 100 = \text{percentage of correct responses.}$$

FIGURE 4–2 Choosing measurement conversation for permanent product data.

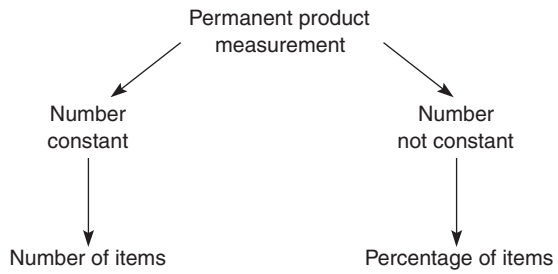


FIGURE 4–3 ***Transferring permanent product data to a graph.***

Student	Catherine		
Behavior	writing paragraphs of 30 words given title and topic sentence		
Date	**Number of Words**	**Date**	**Number of Words**
1 3/16	16	6 3/27	18
2 3/18	24	7 3/30	24
3 3/20	20	8 4/2	20
4 3/23	20	9 4/4	24
5 3/25	22	10 4/7	25

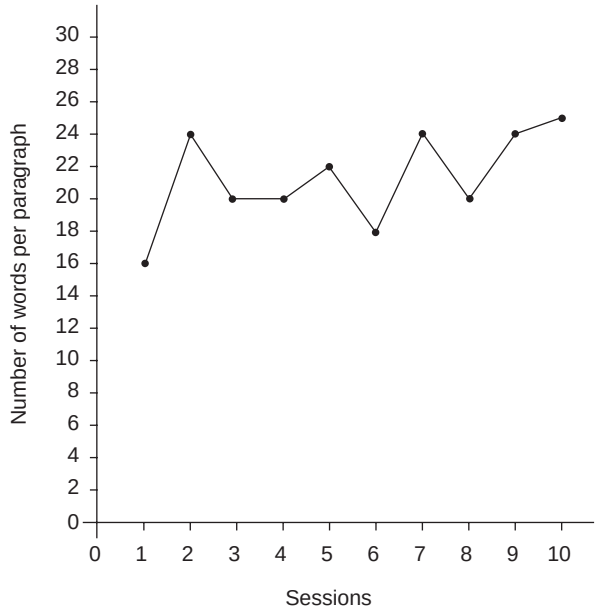


Figure 4–3 is used to record Catherine's performance of paragraph writing. Recorded on the data sheet is the number of words Catherine wrote in each paragraph. Below the data sheet is a simple line graph on which the data have been plotted.

Transferring Event Data to a Graph

Graphing event data.

Event data may be reported as (a) the number of occurrences of a behavior if the amount of time is consistent across sessions, as in the number of times the student was out of seat during the 40-minute math group; (b) the number correct or a percentage if there are a consistent number of opportunities to respond, as in how many sight words out of 10 the student recognized; (c) a percentage correct if the number of opportunities to respond varies, as in the number of times a student complies with a teacher's instruction when the number of instructions varies from session to session. Figure 4–4 is used to record Michael's talking out during a class activity scheduled from 10:20 a.m. to 11:00 a.m. daily. The teacher tallied the number of times Michael called out without raising his hand. Below the data sheet is a graph representing the data. Figure 4–5 is used to record Tasha's recognition of her list of 10 sight words. Below the data sheet, her performance is transferred onto a graph in two ways—as the number of words read correctly and as the percentage of words read correctly.

FIGURE 4–4 ***Transferring event data to a graph.***

Student	Michael
Behavior	calling out without raising hand
Observation Period	10:20 a.m.– 11:00 a.m. (whole-class activity)

Days	Instances	Total
1 Monday	///	3
2 Tuesday	/	1
3 Wednesday	~~////~~ //	7
4 Thursday	///	3
5 Friday	//	2
6 Monday	~~////~~	5
7 Tuesday	////	4
8 Wednesday	////	4
9 Thursday	~~////~~ //	7

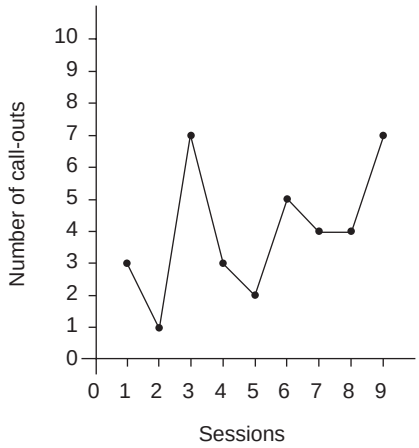
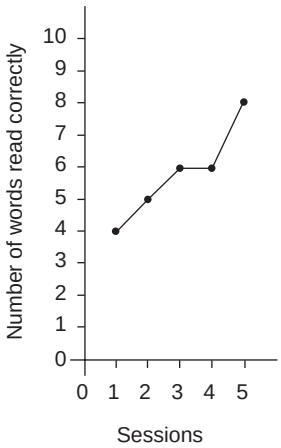
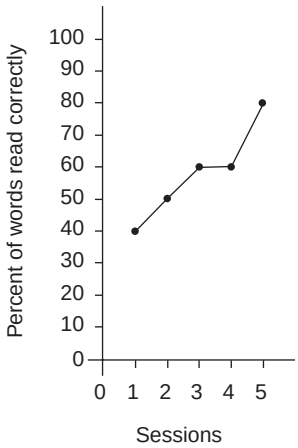


FIGURE 4–5 ***Transferring event data to a graph.***

Student	Tasha
Behavior	sight word reading

	Monday	Tuesday	Wednesday	Thursday	Friday
mother	✓	✓	✓	✓	✓
father	✓	✓	✓	✓	
sister					
brother		✓	✓	✓	✓
school	✓	✓	✓	✓	✓
grocery					
hospital			✓	✓	✓
police	✓	✓	✓	✓	✓
church					✓
station					✓
Total Correct	4	5	6	6	7





Graphing rate data.

Transferring Rate Data to a Graph

Conversion to rate data is required when the teacher is concerned with both accuracy and speed. Rate data reflect fluency of performance and allow judgments about the development of proficiency. If the time allowed for the response(s) is the same across all sessions, simply reporting frequency is all that is necessary. Such is the case when each day the student has 20 minutes to complete a set of 14 math problems. If, however, the time allocated for responding varies from session to session, rate must be calculated so that the data can be compared. Computation of rate is reviewed in Figure 4–6.

Figure 4–7 is used to record Steven's performance during vocational training at the local Red Cross office. Steven was learning to assemble packets of materials used during the blood drive. Because this was vocational training, his teacher was interested in both the number of packets he completed and how long it took him. She was interested in the rate at which Steven assembled packets. Below the data sheet is a graph displaying Steven's rate per minute of assembling packets.

FIGURE 4–6 Computing rate of correct or error responses.

Computing Rate

A rate of correct responding is computed by dividing the number of correct responses by the time taken for responding:

$$\text{Correct rate} = \frac{\text{Number correct}}{\text{Time}}$$

For example, if on Monday Kevin completed 15 problems correctly in 30 minutes, his rate of problems correct would be 0.5 per minute.

$$\frac{\text{15 problems correct}}{\text{30 minutes}} = \text{0.5 problems correct per minute}$$

If on Tuesday he completed 20 problems correctly in 45 minutes, his rate per minute would be 0.44.

$$\frac{\text{20 problems correct}}{\text{45 minutes}} = \text{0.44 problems correct per minute}$$

If Kevin s teacher had simply recorded that Kevin completed 15 problems on Monday and 20 problems on Tuesday, the teacher might think that Kevin's math was improving. In reality, though the number of math problems increased, the rate decreased, and Kevin did not do as well on Tuesday as on Monday.

Computing a rate of error may be done by dividing the number of errors by the time. For example:

$$\text{Session 1: } \frac{\text{12 spelling errors}}{\text{20 minutes}} = \text{0.6 errors per minute}$$

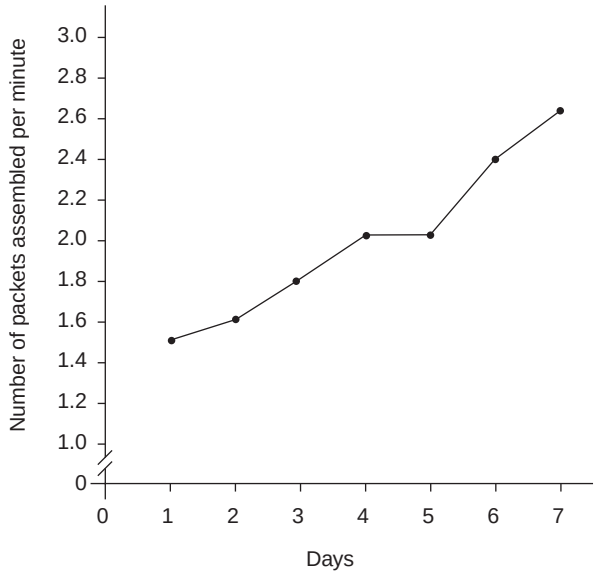
$$\text{Session 2: } \frac{\text{10 spelling errors}}{\text{30 minutes}} = \text{0.33 errors per minute}$$

These rate computations provide the teacher with the numbers of correct or incorrect responses per minute (or second or hour).

FIGURE 4–7 Transferring rate data to a graph.

Student	Steven
Behavior	packet assembly
Observation Period	vocational training at Red Cross

Day	Number Completed	Amount of Time	Rate per Minute
4/16 Monday	45	30´	1.5
4/18 Wednesday	40	25´	1.6
4/20 Friday	45	25´	1.8
4/24 Tuesday	40	20´	2.0
4/26 Thursday	50	25´	2.0
4/30 Monday	48	20´	2.4
5/2 Wednesday	54	20´	2.7



Transferring Interval and Time Sampling Data to a Graph

Graphing interval and time sampling data.

Interval and time sampling data are reported as the number or percent of total observed intervals during which the behavior occurs. They are usually reported as percentages.

Figure 4–8 shows Omar's out-of-seat behavior during the first 6 minutes of center time. The teacher recorded interval data. She constructed the data sheet to show the 6 minutes divided into 20-second intervals and made an X if Omar was out of his seat at any time during the interval. Below the data sheet, the data are transferred onto one graph indicating the number of intervals during which out-of-seat behavior was observed and another graph indicating the percentage of intervals during which Omar was out of his seat at some time.

FIGURE 4–8 ***Transferring interval data to a graph.***

Student: Omar
Behavior: out-of-seat (X = out of seat)
Observation Period: 6 minutes (first 6 minutes of center time)

	20"	40"	60"	20"	40"	60"	20"	40"	60"	20"	40"	60"	20"	40"	60"	20"	40"	60"
Mon	—	X	X	X	—	—	X	—	—	—	—	—	X	X	X	—	X	X
Tues	X	X	X	X	—	X	—	—	—	—	X	—	—	—	—	—	X	—
Wed	—	—	—	—	X	X	—	—	—	—	—	—	—	X	X	—	—	—
Thur	X	—	X	X	—	X	X	—	—	—	X	X	—	X	—	—	—	X
Fri	—	X	—	—	—	X	—	X	—	X	X	—	—	X	X	X	—	X

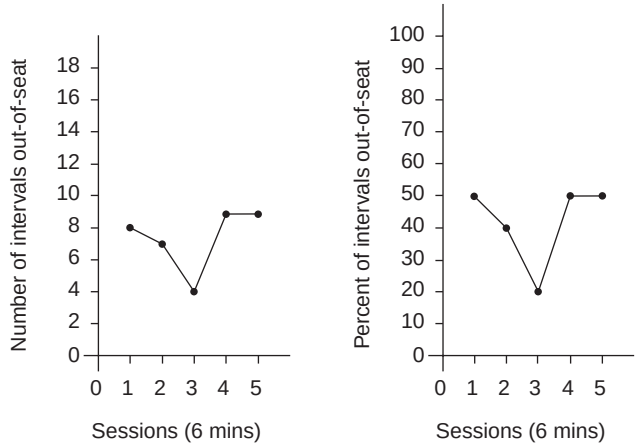


Figure 4–9 presents the data sheet on which the teacher recorded whether Leann was engaged in self-talk or peer-directed talk during 20-minute play periods. The teacher chose to use time sampling; she recorded the type of talk in which Leann was engaged at the end of each interval. Below the data sheet, the data are transferred onto one graph indicating number of intervals and another indicating percent of intervals. Note that each graph uses a different identification system to indicate the behavior associated with a data path.

FIGURE 4–9 Transferring time sampling data to a graph.

Student: Leann
Behavior: self-talk vs. other directed during play (S—self, O—other)
Observation Period: 20 minutes play period (varies each day)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mon	—	S	S	S	—	O	O	—	—	S	S	S	S	S	—	O	O	—	S	S
Tues	—	—	—	—	—	S	S	S	S	—	—	—	—	—	—	—	—	—	—	—
Wed	—	—	O	—	—	—	S	S	S	—	—	S	S	S	—	S	S	—	—	O
Thur	O	S	—	—	—	S	S	S	S	—	O	O	O	S	—	—	O	O	S	S
Fri	—	S	S	—	—	S	S	S	O	S	S	S	O	—	—	—	S	S	O	O

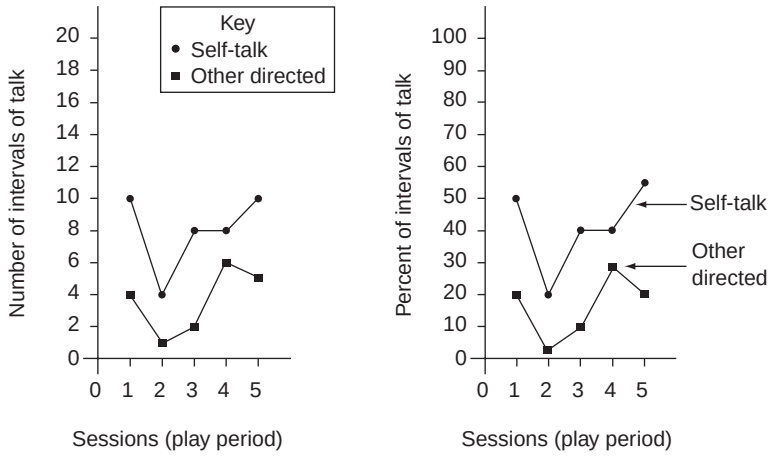
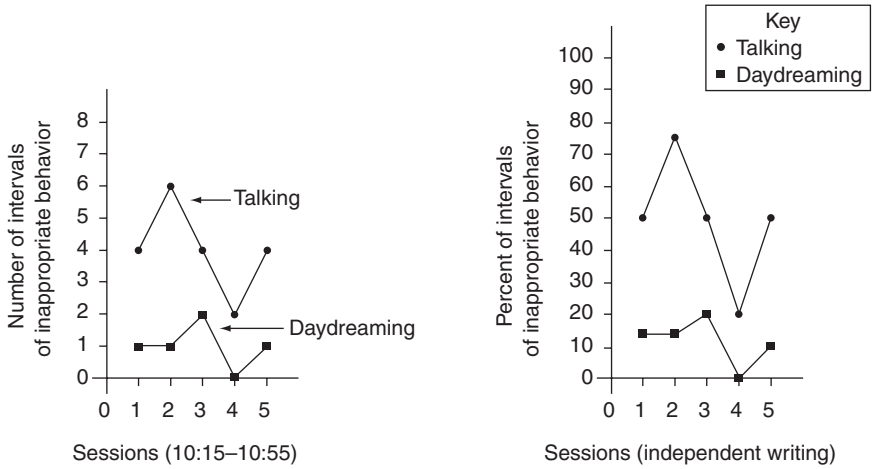


Figure 4–10 presents a differently arranged data sheet for time sampling. On this data sheet the teacher indicated the nature of Kosh's off-task behavior during the time allocated for independent writing. Each of the two types of off-task behavior would have been properly operationalized before data collection. Below the data sheet are two ways of graphing these data and two ways of keying a graph.

FIGURE 4–10 Transferring time sampling data to a graph.

Student	Kosh			
Behavior	off-task (T—talking, D—daydreaming)			
Observation Period	independent writing (10:15–10:55)			
Monday	T 5′	T 10′	— 15′	— 20′
	— 25′	D 30′	T 35′	T 40′
Tuesday	T 5′	T 10′	D 15′	T 20′
	T 25′	— 30′	T 35′	T 40′
Wednesday	T 5′	T 10′	— 15′	— 20′
	D 25′	D 30′	T 35′	T 40′
Thursday	— 5′	— 10′	T 15′	T 20′
	— 25′	— 30′	— 35′	— 40′
Friday	T 5′	D 10′	T 15′	T 20′
	— 25′	— 30′	— 35′	T 40′



Graphing duration data.

Transferring Duration Data to a Graph

Duration data may be collected and reported either as the number of minutes or seconds it takes a student to complete a behavior or as how much of a specified period of time a student spent engaging in a particular behavior. A teacher might record, for example, the total amount of time it took a student to finish an assigned task. Another teacher might record how much of a 20-minute science lab a student spent engaged in the lab project. The second example could be reported as the number of minutes of engagement or as the percentage of available time spent engaged in the project.

FIGURE 4–11 Transferring duration data to a graph.

Student Behavior	Casey time spent toileting			
Monday	1	12 minutes		
	2	8 minutes	➜	average = 9 minutes
	3	7 minutes		
Tuesday	1	11 minutes		
	2	16 minutes		
	3	9 minutes	➜	average = 12 minutes
Wednesday	1	15 minutes		
	2	10 minutes	➜	average = 11 minutes
	3	8 minutes		
Thursday	1	14 minutes		
	2	10 minutes		
	3	12 minutes	➜	average = 12 minutes
Friday	1	9 minutes		
	2	11 minutes		
	3	10 minutes	➜	average = 10 minutes

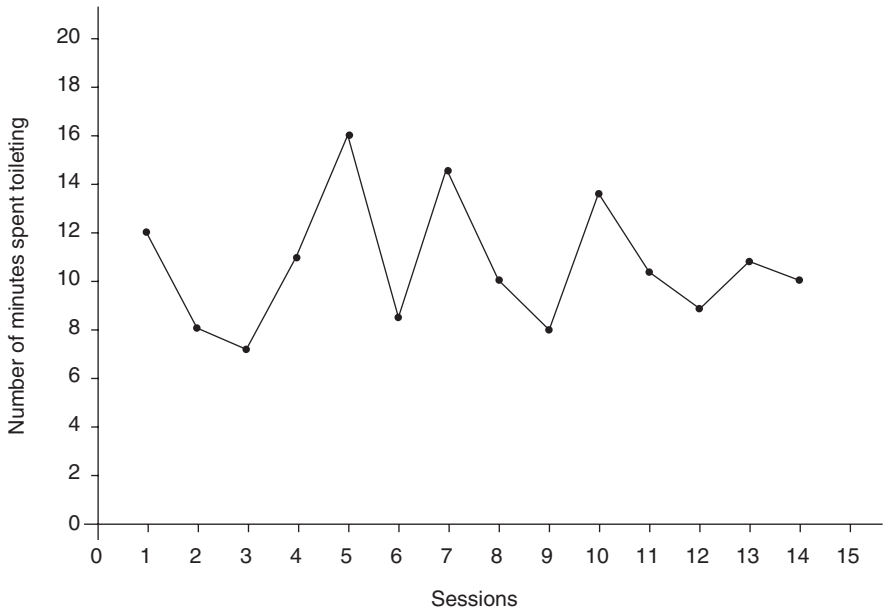


Figure 4–11 shows how long it takes Casey to complete toileting. The graph below the data sheet is drawn to indicate the number of minutes he spent each time he went to the bathroom.

Transferring Latency Data to a Graph

Graphing latency data.

Latency data are reported as the number of minutes or seconds that elapse before a student initiates a behavior following a request for the behavior to be performed or for a natural occasion for its performance to occur (for example, answering a ringing telephone).

FIGURE 4–12
Transferring latency data to a graph.

Student	DuShawn
Behavior	delay beginning of morning paragraph writing
Observation Period	each morning—8:45 a.m.

Day	Number of Minutes	Comments
Monday	6	pencil sharpening
Tuesday	5	roaming
Wednesday	6	pencil sharp
Thursday	2	chat
Friday	4	chat
Monday	5	pencil sharp
Tuesday	7	pencil sharp
Wednesday	5	coat-pencil
Thursday	4	roaming
Friday	5	pencil sharp

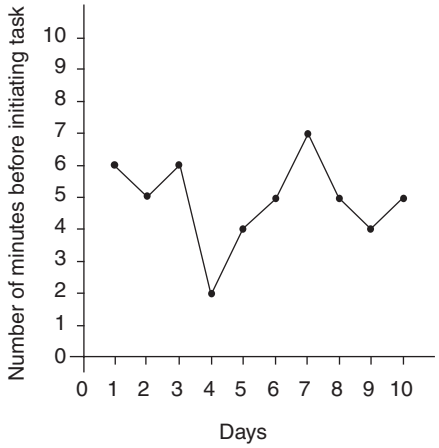


Figure 4–12 shows DuShawn's latency in beginning his daily journal-writing exercise. After giving the class the instruction to begin, the teacher recorded how many minutes elapsed before DuShawn began the task and noted what he was doing before beginning the task. Below the data sheet is a graph that represents these data.

ADDITIONAL GRAPHING CONVENTIONS

More complex graphs than the ones in this chapter will be illustrated in Chapter 6. Some additional conventions that will help you understand them are described in Figure 4–13.

Conditions are phases of an intervention during which different approaches or techniques are used. A teacher who wanted to reduce the occurrence of an inappropriate behavior might first record the current level of the behavior for several sessions or days—called **baseline data**—and then he would use some strategy to assist the student to decrease her performance of the behavior—called **intervention.** You need a clear indicator on the graph of which condition is in effect during each session. This is

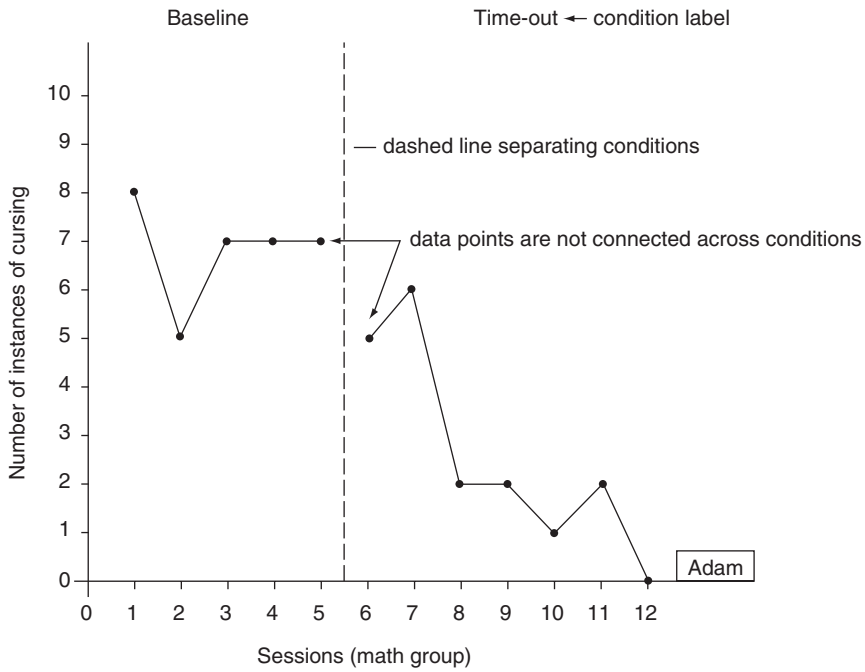


FIGURE 4–13 ***Conventions for conditions represented on a graph.***

provided by drawing a vertical dashed line from the top to the bottom of the graph. This line is drawn between the lines on the graph paper between the last session of one condition and the first session of the next. For example, if baseline occurred for five sessions and intervention began on the sixth session, the condition line would be drawn between sessions 5 and 6 (represented on the abscissa). Data points are not connected across conditions. To identify what procedure is represented, a brief, descriptive condition label is placed above the data path for each condition, centered between the vertical dashed lines. For example, if a teacher is using a time-out procedure to reduce a student's cursing, he would print "time-out" centered above the part of the graph on which these data are placed.

CUMULATIVE GRAPHS

On a simple line graph, data points are plotted at the appropriate intersections without regard to performance during the previous session. On a **cumulative graph,** the number of occurrences observed in a session is graphed after being added to the number of occurrences plotted for the previous session. The occurrences recorded for each session include those of all previous sessions. A cumulative graph presents an additive view of a behavior across sessions, providing a count of the total number of responses. The hypothetical graphs in Figure 4–14 show the same raw data plotted on a line graph and on a cumulative graph.

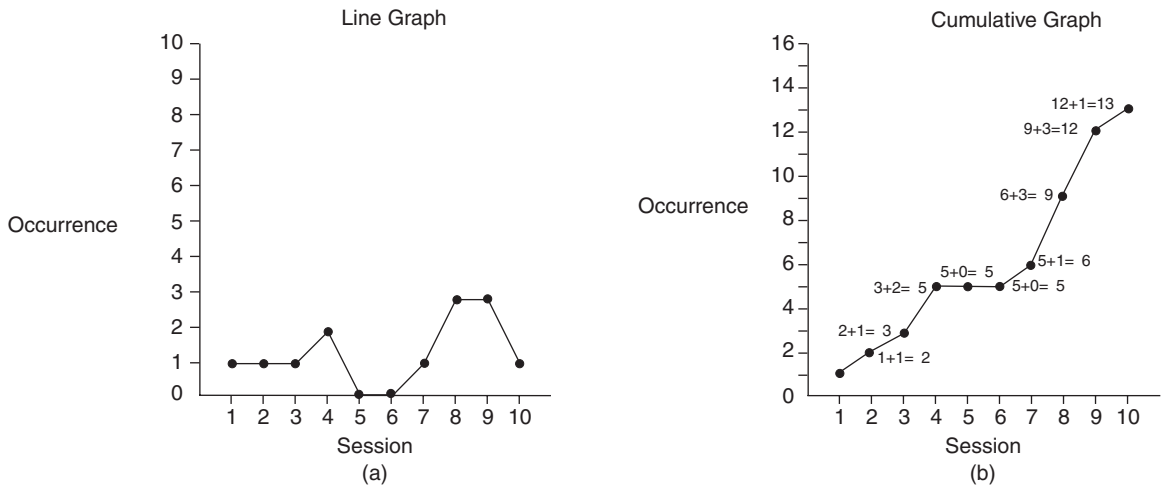
Bar graphs are clearer for young students.

Cumulative graphs always demonstrate an upward curve if any behavior at all is being recorded (Ferster, Culbertson, & Boren, 1974). This form of data illustration provides a continuous line with a slope that indicates the rate of responding. A steep

FIGURE 4–14 Comparison of plotting data points on line and cumulative graphs.

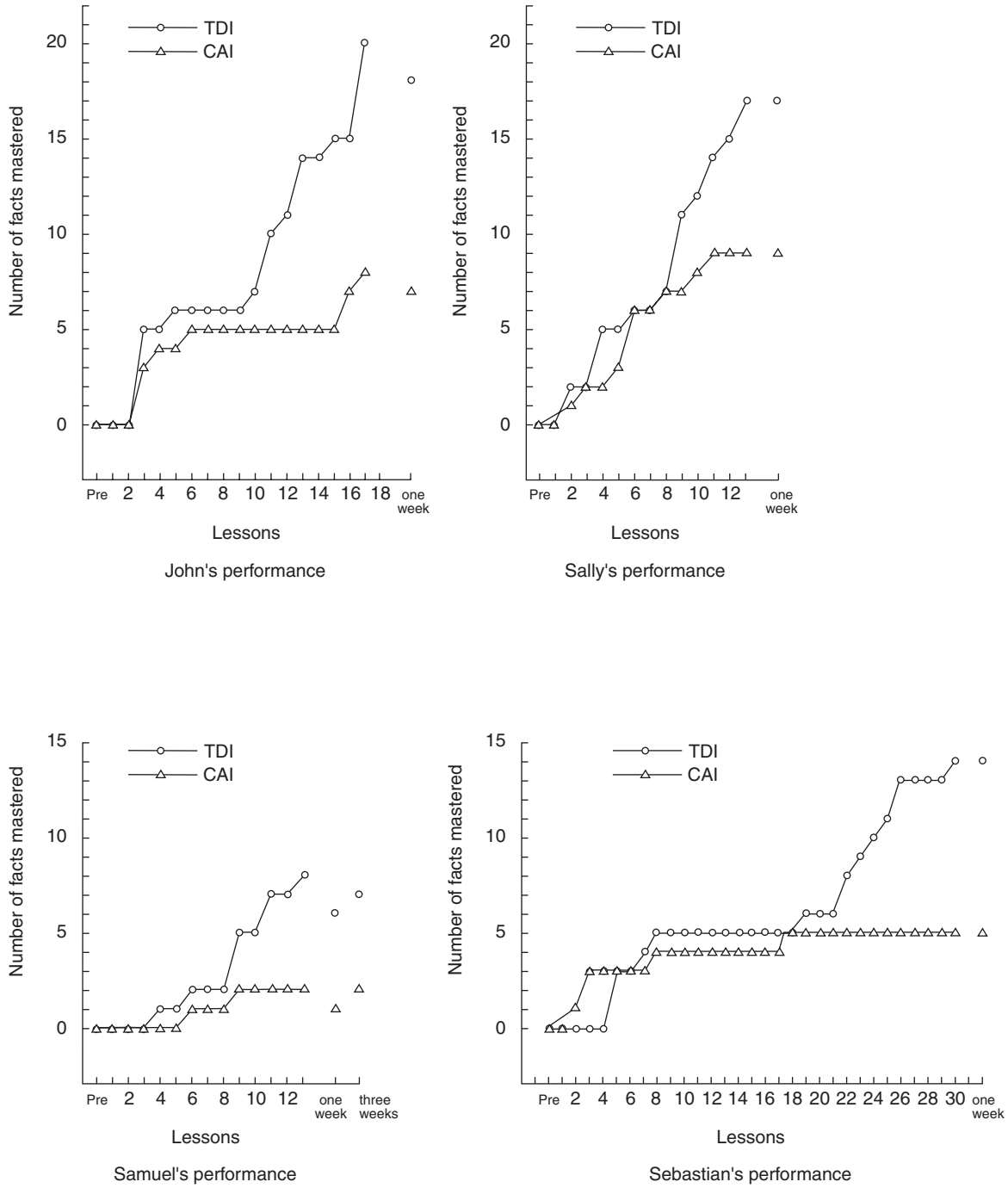
Raw Data

Session	Occurrences
1	1
2	1
3	1
4	2
5	0
6	0
7	1
8	3
9	3
10	1



slope indicates rapid responding, a gradual slope indicates slow responding, and a plateau or straight line indicates no responding. Each of these slope variations can be seen in the results of the study by Wilson, Majsterek, and Simmons (1996). These authors compared the use of teacher-directed instruction (TDI) with computer-assisted instruction (CAI) on the acquisition of multiplication facts by elementary students with learning disabilities. Under TDI, flash cards were used, correct responses were verbally praised, and errors were corrected immediately by repeating a three-step instructional sequence (teacher alone, student and teacher together, student alone). Under CAI, a popular computer math software program was used that delivered a brief praise statement for correct responses, indicated errors, and provided a second attempt to answer. If the second attempt was incorrect, the software provided the correct answer and the student was instructed to repeat the answer. All students mastered more facts in the teacher-directed condition. The authors attribute the greater success of TDI to "the flexible and responsive nature of the teacher-directed condition (that) allowed teachers to move at a quicker pace, remediate errors more immediately and differentially, and build fluency through greater opportunities for practice" (p. 389). Figure 4–15 presents the graphed data of the four participating students. The authors chose to design the graph so that readers could see students' cumulative acquisition of multiplication facts. Several examples of the three types of slopes just described can be clearly seen in these graphs. Examine, for example, the steep slope indicating rapid responding

FIGURE 4–15 ***Examples of cumulative graphs.***



on John's graph for lessons 10 through 13 of TDI, the gradual slope indicating slower responding on Samuel's graph for lesson 9 through 13 of TDI, and the plateaus or straight lines indicating that Sebastian mastered no new facts during lessons 18 through 30 under the CAI condition.

You can't ski this slope.

Bar Graphs

A **bar graph,** or histogram, is yet another means of displaying data. Like a line graph, a bar graph has two axes, the abscissa (sessions) and the ordinate (performance). As its name implies, a bar graph uses vertical bars rather than data points and connecting lines to indicate performance levels. Each vertical bar represents one observation period. The height of the bar corresponds with a performance value on the ordinate. A bar graph may be better for displaying data in situations where clear interpretation of the pattern of behavior plotted on a line graph is difficult. Such confusion may result when several data paths are plotted on a single line graph, as when a teacher chooses to include data from several students or from multiple behaviors. In such cases, plotted lines may overlap or appear extremely close together because data points fall at the same intersections. Figure 4–16 offers an example of the same data plotted on a line graph and on a bar graph. The bar graph is plainly much clearer. A classroom teacher might use a bar graph to display daily the number of correct responses from each member of a small group.

Another use of the bar graph is to summarize student performance data (Tawney & Gast, 1984). This may be done for a single task, such as the mean number of science tasks completed across multiple students (Figure 4–17), or to summarize a single student's performance across multiple tasks (Figure 4–18).

FIGURE 4–16 ***Comparing line and bar graphs.***

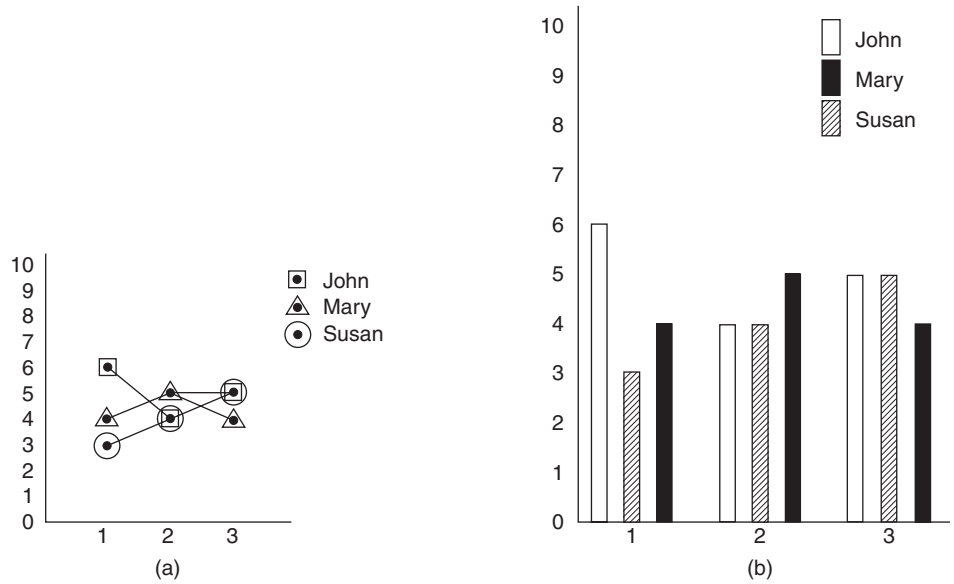


FIGURE 4–17 ***Bar graph of summary of task performance of multiple students.***

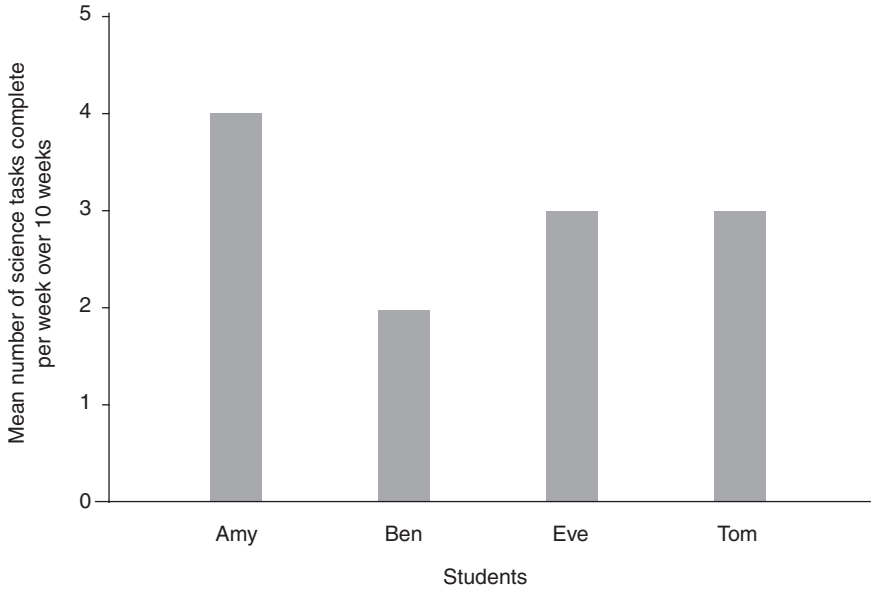
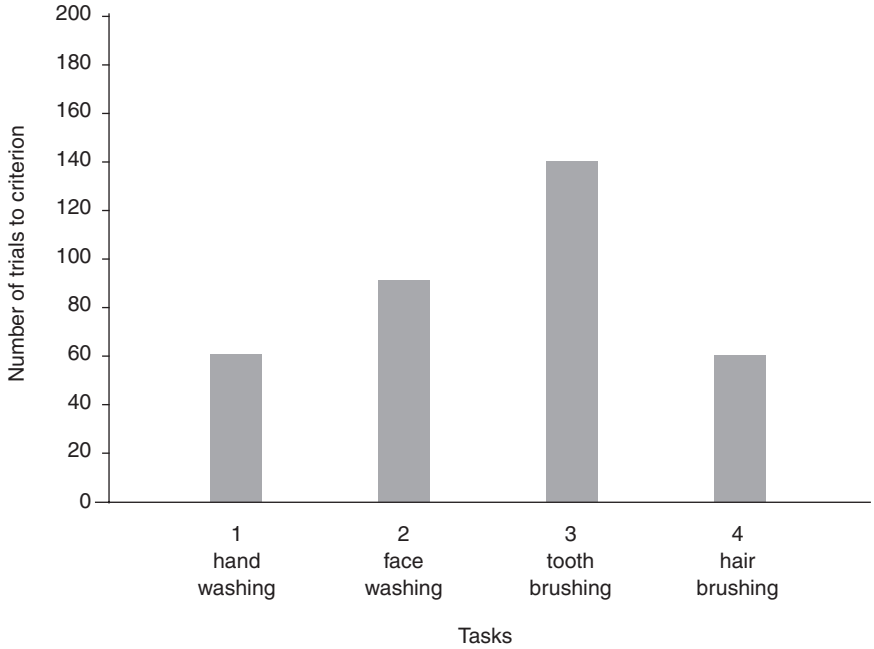


FIGURE 4–18 ***Bar graph of summary of student performance across tasks.***



SUMMARY

We discussed the basic reasons for graphing data, including monitoring of student performance, formative and summative evaluation, and graphs as a tool for communicating among educators, students, and parents. We described three basic methods of graphing data: line graphs, cumulative graphs, and bar graphs. We outlined the basic conventions for drawing line graphs and provided examples of transferring various types of observational data from data sheets to graphs.

KEY TERMS

abscissa
ordinate
conditions
baseline data
intervention
cumulative graph
bar graph

DISCUSSION QUESTIONS

1. How can graphs be used to (a) summarize and report student progress and (b) communicate with students, parents, and related service professionals?
2. What is the difference between frequency data and rate data? When is using each most appropriate?
3. Develop a data sheet and a graph (with hypothetical data) for the following: How much of a 20-minute science lab does Cary spend on the assigned project rather than being off task?
4. Are line graphs or bar graphs a better mechanism for monitoring ongoing student performance?
5. How and with whom can graphs be used as a communication tool?

Chapter 5

Single-Subject Designs

Did you know that . . .

- Not all sixth graders should use the same math text?
- Not every behavior change is functionally important?
- A success is truly a success only when you can replicate it?
- Single-subject research is an instructional tool?
- Big changes can happen in small steps?

CHAPTER OUTLINE

Data collection allows teachers to make statements about the direction and magnitude of behavioral changes. Data collection alone, however, does not provide sufficient information to indicate a functional relation between an intervention and the behavior in question. To make assumptions about functional relations, data collection must be conducted within certain formats, or designs. A design is a systematic pattern for collecting data that enables the collector to make confident statements about the relation between interventions and behaviors.

In this chapter we will describe a number of experimental designs used in applied behavior analysis that enable teachers and researchers to determine relations between interventions and behavior change. Each design has a particular graphic format. The

various formats are what allow visual inspection and analysis of the data. Graphs can be made with paper, a ruler, and a pencil; in most cases, however, the result will look like part of a middle school project. Such graphs do not present the image a teacher wants for a meeting with parents and other professionals gathered to evaluate the progress of student learning. For a professional appearance, standardized graphs can be made using computer software such as Microsoft Excel. To assist in creating graphs with computers, *Graphing in Excel for the Classroom: A Step-by-Step Approach* by Cihak, Alberto, Troutman, and Flores (2005) provides step-by-step instructions, with picture prompts, for creating graphs for classroom use and for publication.

Teachers who can read and understand experimental research reported in professional journals can remain up to date on innovative techniques and procedures. Learning about these designs may also encourage them to become teacher-researchers who can systematically evaluate their own instruction and share their results with others. The ability to conduct classroom-based research will increase teachers' confidence, effectiveness, and credibility.

In this chapter, research applications, taken from professional journals, accompany the description of each design. Each design is also applied to a classroom problem to demonstrate the utility of applied behavior analysis designs in the classroom.

Variables and Functional Relations

We will define some terms basic to experimental investigation before discussing specific designs. The term variable is used to refer to any number of factors involved in research. These may include attributes of individuals being studied (age, test scores), conditions associated with the setting in which the study is done (number of students, noise level), or the nature of an intervention, which might be an instructional strategy (direct instruction phonics, cooperative learning), instructional material (counting chips, computer), or behavior management technique (tokens, self-recording). In research, the goal is to control for the presence or absence of variables that may affect outcomes. An unforeseen or uncontrolled variable (illness, for example) is referred to as a *confounding variable*. If a teacher is using a new program to teach a student long division, and the student's father coincidentally begins reviewing long division with the student for an hour every evening, it will be impossible to determine whether the teacher's variable (new math program) or the uncontrolled variable (home instruction) was responsible for the student's learning long division. With experimental designs researchers can control for many confounding variables.

Experimental design differentiates between two types of variables: dependent and independent. The term **dependent variable** refers to the behavior targeted for change. The term **independent variable** refers to the intervention being used to change behavior. In the following sentences, the independent variable is *italicized* and the dependent variable appears in (parentheses).

Following a student's (oral reading), the teacher provided *corrective feedback*.

Picture prompts are used when the student is engaged in (buying groceries).

Coworker modeling is provided when the student is (shelving books) in the library.

Verbal praise is contingently presented when the student (remains on task for 15 minutes).

Contingent upon (a temper tantrum), the student is placed in *time-out*.

For each (math problem completed correctly) the student earned *one token*.

Single-subject experimental designs allow researchers to assess cause and effect between an independent variable and a dependent variable. When changes in a dependent variable are *replicated* each time the same independent variable is implemented, a **functional relation** is said to exist. Single-subject experimental designs provide the framework within which to test for these replications of effect. When interventions and their outcomes are replicated, the teacher and researcher can have confidence that the behavior changed as a function of the intervention because with each replication only the independent variable is changed or manipulated. Repeated manipulation allows the teacher-researcher to rule out confounding variables as agents of the behavior change. Additionally, the demonstration of a functional relation is evidence of experimental control (Kennedy, 2005). This experimental control adds confidence that some extraneous or confounding variable(s) was not the cause of the effect.

Basic Categories of Designs

Group designs versus single-subject designs.

A research design is a format that structures the manner in which questions are asked and data are collected and analyzed. Two categories of research designs are **group designs** and **single-subject designs.** Each provides a plan and a means for demonstrating the effectiveness of an intervention on a behavior. As indicated by their names, group designs focus on questions and data related to groups of individuals, whereas single-subject designs focus on questions and data related to a particular individual.

Group designs are used to evaluate the effects of an intervention on a behavior of a whole population (for example, all the second graders of a school district or of a school building) or of a representative sample of a population. In order to determine the effectiveness of an intervention, the population (or a randomly selected sample) is, also randomly, divided into two groups: an experimental group and a control group. (It is this random selection and division that allows generalization from a sample to an entire population.) Members of the experimental group receive the intervention. This provides multiple replications of the effect of the intervention. Members of the control group do not receive the intervention. Measurements of the behavior (averages in performance) are made before intervention and at the conclusion of intervention for each group. Average changes in behavior of the two populations are compared subsequent to the intervention. This comparison is made through use of statistical procedures, the purpose of which is: (a) to verify a difference in the change in average scores between the two groups, (b) to verify that the difference is significant and therefore possibly worthy of being acted upon, and (c) to verify that the differences between groups are more likely the result of the intervention than of chance or some unknown source.

For example, the curriculum committee of Fulton County Public Schools is considering changing their sixth-grade math text. Currently they use the text by Jones and Jones. The committee randomly selects 200 students from among all the sixth graders in the district. These students are then randomly assigned to either the experimental group (100 students) or the control group (100 students). During the first week of the school year all 200 students are tested on sixth-grade math objectives. Then during the school year the experimental group receives instruction using the Smith and Smith math text, while the control group continues to receive instruction using the Jones and Jones math text. At the end of the school year each group is tested again on sixth-grade math objectives. The average gain in performance (number of objectives met) for the groups is compared. This is done to determine: (a) if there is no difference in gain in scores of the two groups, (b) if there is a difference, whether this difference is significant, and (c) whether it is reasonable to assume that a greater or lesser gain in score by the experimental group is due to using the Smith and Smith text.

Applied behavior analysts prefer multiple measurements of behavior in order to provide a detailed picture of the behavior before and during the course of intervention. They also prefer to record information specific to individuals rather than information about average performance of groups. Examining average performance may obscure important information, as illustrated in the following anecdote.

Ms. Witherspoon Orders Reading Books

Ms. Witherspoon, a third-grade teacher, was urged by her principal to order new reading books at the beginning of the school year. Being unfamiliar with her class, Ms. Witherspoon decided to use a reading test to determine which books to order. She administered the test and averaged the scores to determine the most appropriate reader. She came up with an exact average of third grade, first month, and ordered 30 readers on that level.

When the books arrived, she found that the reader was much too hard for some of her students and much too easy for others. Using an average score had concealed the fact that although the class average was third grade, some students were reading at first-grade level and others at sixth-grade level.

Single-Subject Designs

Single-subject designs often compare the effects of different conditions on the same individual.

Applied behavior analysis researchers prefer to use single-subject designs. Single-subject designs provide structures for evaluating the performance of individuals rather than groups. Whereas group designs identify the effects of variables on the average performance of large numbers of students, single-subject designs identify the effects of variables on a specific behavior of a particular student. These designs monitor the performance of individuals during manipulation of the independent variable(s). Certain techniques, described later in the chapter, are used to verify that changes in the dependent variable result from experimental manipulations and not from chance, coincidence, or confounding variables.

Single-subject designs require repeated measures of the dependent variable. The performance of the individual whose behavior is being monitored is recorded weekly, daily, or even more frequently over an extended period of time. The individual's performance can then be compared under different experimental conditions, or manipulations of the independent variable. Each individual is compared only to himself or herself, though the intervention may be replicated with several other individuals within the same design. Single-subject research emphasizes clinical significance for an individual rather than statistical significance among groups. If an intervention results in an observable, measurable improvement in functioning, often referred to as **enhanced functioning,** the results of the experiment are considered to have clinical significance.

Certain components are common to all single-subject designs. These include a measure of baseline performance and at least one measure of performance under an intervention condition. Single-subject research designs require at least one replication of the use of the intervention within the design. This replication allows for the assumption of a functional relation.

Applied behavior analysts do not assume generality of research results based on a single successful intervention. When a functional relation is established between an independent variable (intervention) and a dependent variable (behavior) for one individual, repeated studies of the same intervention are conducted using different individuals and different dependent variables. The more frequently an intervention proves effective, the more confidence is gained about the generality of the results of the intervention. That systematic teacher praise increases one student's rate of doing math problems may not be a convincing argument for the use of praise. Documentation that such praise

increased production of not only math problems but also other academic and social behaviors with numerous students is more convincing. Using systematic replication, applied behavior analysts gradually identify procedures and techniques effective with many students. Others can then adopt these procedures and techniques with considerable confidence that they will work.

Sidman (1960) suggested that it would be an error to view single-subject research as simply a microcosm of group research. Repeated measures of a dependent variable when the independent variable is applied and removed demonstrate a continuity of cause and effect and the relation of one data point to another that would not be seen when comparing the effect of the independent variable across separate groups. He contends that individual and group curves do not provide the same information, "for the two types of data represent, in a very real sense, two different subject matters" (p. 54).

Baseline Measures

The first phase of single-subject design involves the collection and recording of baseline data. **Baseline data** are measures of the level of behavior (the dependent variable) as it occurs naturally, before intervention. Kazdin (1982, 1998) stated that baseline data serve two functions. First, baseline data serve a *descriptive function*. These data describe the existing level of student performance. When data points are graphed, they provide a picture of the student's behavior—his current ability to solve multiplication problems or his current rate of talk-outs. This objective record can assist the teacher in verifying the existence and extent of the behavior deficit (lack of ability to do multiplication) or behavior excess (talking out).

Second, baseline data serve a *predictive function*. "The baseline data serve as the basis for predicting the level of performance for the immediate future if the intervention is not provided" (Kazdin, 1982, p. 105). To evaluate the success of an intervention (the independent variable), the teacher must know what student performance was like before the intervention. Baseline data serve a purpose similar to that of a pretest. "The predication is achieved by projecting or extrapolating into the future a continuation of baseline performance" (p. 105). It is against this projection that the effect of an intervention is judged.

Baselines should be stable.

The baseline phase continues for several sessions before the intervention phase begins. In most instances, at least five baseline data points are collected and plotted. The extent of baseline data collection is affected by certain characteristics of these data points.

See Chapter 2 for suggestions on writing operational definitions.

Because baseline data are to be used to judge the effectiveness of the teacher's intervention, it is important that the baseline be *stable*, providing a representative sample of the natural occurrence of the behavior. The *stability* of a baseline is assessed by two characteristics: variability of the data points and trends in the data points. *Variability of data* refers to fluctuations in the student's performance. "As a general rule, the greater the variability in the data, the more difficult it is to draw conclusions about the effects of the intervention" (Kazdin, 1982, p. 109) and to make projections about future performance. When baselines are unstable, the first thing to examine is the definition of the target behavior. A lack of stability in the baseline may suggest that the operational definition of the target behavior is not sufficiently descriptive to allow for accurate and consistent recording or that the data collector is not being consistent in the procedure used for data collection. In laboratory settings, other sources of variability can often be identified and controlled. In classrooms, attempts to control variability are desirable if the sources of variability can be identified—for example, if fluctuations are caused by inconsistent delivery of medication. In cases of temporary fluctuations caused by such unusual events as a fight or a problem at home, the teacher may just wait for the fluctuation to pass. However, in classrooms, unlike laboratories, "variability is an

Session	Data Points
1	14
2	10
3	20
4	16
5	11

Baseline mean (arithmetic average) = 14.2 = 14
50% of mean = 7
Acceptable range of data points = 7 − 21 (14 ± 7)
This baseline is stable because no data point varies more than 50% from the mean.

FIGURE 5–1
Computing baseline stability.

unavoidable fact of life," and in such settings there are seldom "the facilities or time that would be required to eliminate variability" (Sidman, 1960, p. 193).

Where variables can be rigorously controlled, a research-oriented criterion for the existence of variability would be data points within a 5% range of variability (Sidman, 1960). A therapeutic criterion of 20% has been suggested (Repp, 1983). However, in classrooms where pure research concerns might be less important than rapid modification of the behavior, we suggest a more lenient parameter of 50% variability. If variability exceeds 50%, statistical techniques for performance comparisons must be used (Barlow & Hersen, 1984). A baseline may be considered stable if no data point of the baseline varies more than 50% from the mean, or average, of the baseline. Figure 5–1 illustrates a procedure for computing the stability of a baseline based on this criterion.

A **trend** in the data refers to an indication of a distinctive direction in the performance of the behavior. A trend is defined as three successive data points in the same direction (Barlow & Hersen, 1984). A baseline may show no trend, an increasing trend, or a decreasing trend. Figures 5–2 and 5–3 illustrate two types of trends—increasing and decreasing.

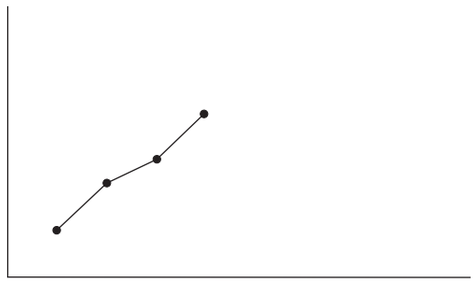


FIGURE 5–2
Increasing trend (ascending baseline).

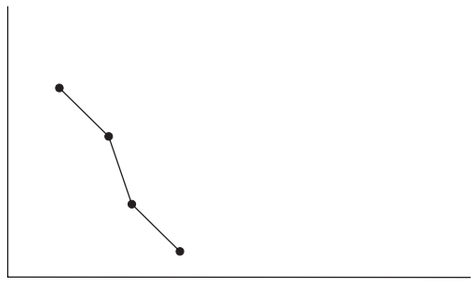


FIGURE 5–3
Decreasing trend (descending baseline).

An *ascending baseline* denotes an increasing trend. Teachers should initiate intervention on an ascending baseline only if the objective is to decrease the behavior. Because the behavior is already increasing, the effects of an intervention designed to increase behavior will be obscured by the baseline trend.

Take baseline trends into account before intervention.

A *descending baseline* includes at least three data points that show a distinctive decreasing direction or trend in the behavior. Teachers should initiate intervention on a descending baseline only if the objective is to increase the behavior.

Intervention Measures

The second component of any single-subject design is a series of repeated measures of the subject's performance under a treatment or intervention condition. The independent variable (treatment or intervention) is introduced, and its effects on the dependent variable (the student's performance) are measured and recorded. Trends in treatment data indicate the effectiveness of the treatment and provide the teacher or researcher with guidance in determining the need for changes in intervention procedures.

Experimental Control

Professor Grundy encounters a confounding variable when he again visits Miss Harper in this chapter.

Experimental control refers to the researcher's efforts to ensure that changes in the dependent variable are in fact related to manipulations of the independent variable—that a functional relation exists. The researcher wants to eliminate to the greatest extent possible the chance that other, confounding variables are responsible for changes in the behavior. Confounding variables are those environmental events or conditions that are not controlled by the researcher but may affect behavior. For example, if a teacher institutes a behavioral system for reducing disruptive behavior in a class after the three most disruptive students have moved away, she really cannot be sure that the new system is responsible for the lower levels of disruption. Removal of the three students is a confounding variable.

The designs discussed in this chapter provide varying degrees of experimental control. Some, called here *teaching designs*, do not permit confident assumption of a functional relation. They may, however, provide sufficient indication of behavior change for everyday classroom use, particularly if the teacher remains alert to the possibility of confounding variables. Other designs, called *research designs*, provide for much tighter experimental control and allow the teacher or researcher to presume a functional relation. Researchers usually demonstrate experimental control by repeating an intervention several times and observing its effect on the dependent variable each time it is repeated. Research designs may be used in classrooms when a teacher is particularly concerned about possible confounding variables and wants to be sure that intervention has had the desired effect on behavior. The teacher who is interested in publishing or otherwise sharing with other professionals the results of an intervention would also use a research design if at all possible. In the following sections on specific research designs, both a research and a classroom application are described for each design whenever possible.

AB Design

The AB design is a teaching design.

The **AB design** is the basic single-subject design. Each of the more complex designs is actually an expansion of this simple one. The designation *AB* refers to the two phases of the design: the A, or baseline, phase and the B, or intervention, phase. During the A phase, baseline data are collected and recorded. Once a stable baseline has been established, the intervention is introduced, and the B phase begins. In this phase, intervention data are collected and recorded. The teacher can evaluate increases or decreases in the amount, rate, percentage, or duration of the target behavior during the intervention phase and compare them with the baseline phase. Using this information

to make inferences about the effectiveness of the intervention, the teacher can make decisions about continuing, changing, or discarding the intervention.

Implementation

Table 5–1 shows data collected using an AB design. The teacher in this instance was concerned about the few correct answers a student gave to questions about a reading assignment. For 5 days, she collected baseline data. She then made 2 minutes of free time contingent on each correct answer and continued to record the number of correct responses. As shown in Table 5–1, the number clearly increased during the intervention phase. The teacher could make a tentative assumption that her intervention was effective.

Graphic Display

Data collected using an AB design are graphed in two phases: A, or baseline, and B, or intervention. A broken vertical line on the graph separates the two phases and data points between phases are not connected. The graph in Figure 5–4 shows a clearer picture of the effectiveness of the intervention than do the data in table form.

TABLE 5–1
Sample data from an AB design.

Baseline Data	
Day	**Number of Correct Responses**
Monday	2
Tuesday	1
Wednesday	0
Thursday	2
Friday	1
Intervention Data	
Day	**Number of Correct Responses**
Monday	6
Tuesday	6
Wednesday	4
Thursday	8
Friday	6

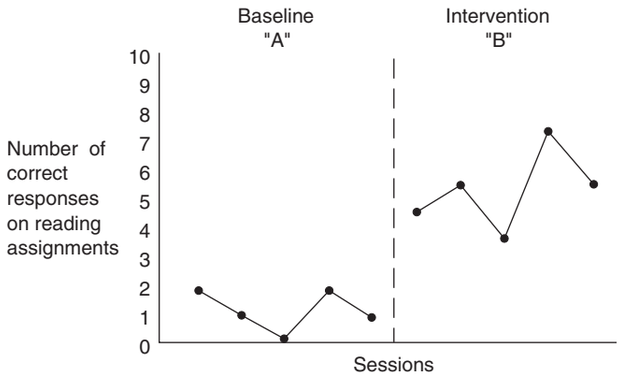


FIGURE 5–4
Graph of AB design data from Table 5–1.

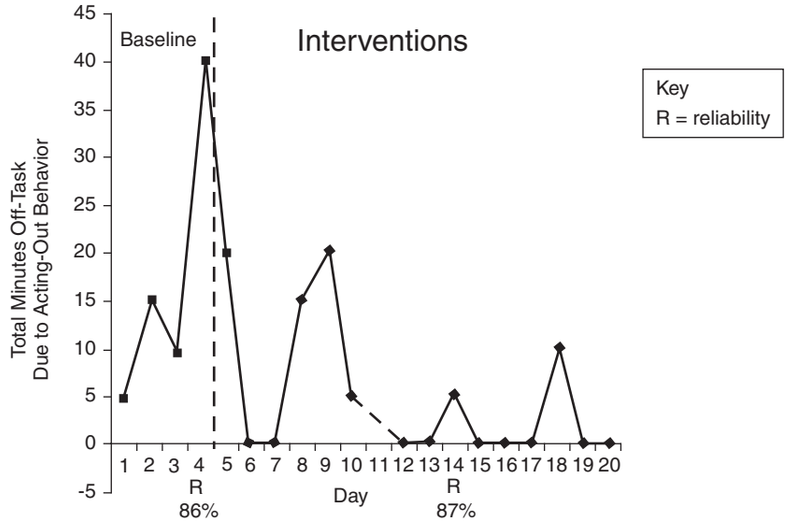


FIGURE 5–5
Use of an AB design.
Note: From "Decreasing Acting-Out Behavior and Increasing Learning," by S. Schoen & J. Nolen, 2004, *Teaching Exceptional Children, 37*, 26–29. Copyright 2004, by The Council for Exceptional Children. Reprinted with permission.

Design Application

The basic AB design is not often found in the research literature because it cannot assess for a functional relation. The design does not provide for the replication within an experiment that establishes a functional relation. Schoen and Nolen (2004) used an AB design to illustrate the results of an intervention designed to reduce the off-task behaviors of a sixth-grade boy with learning disabilities. A self-management checklist was employed with which he assessed his behavior. Figure 5–5 illustrates the decline in the total number of minutes he was off-task from baseline through the intervention phase. One cannot, however, assume a functional relation between the dependent variable (off-task behaviors) and the independent variable (self-management checklist) because the AB design does not provide for repeated manipulation (use and removal) of the independent variable. This study and the suitability of the use of some single-subject methodologies are discussed in the section on action research at the end of this chapter.

The following example demonstrates another use of an AB design in a classroom setting.

Jack Learns to Do His French Homework

Mr. Vogl had difficulty working with Jack, a student in the fourth-period French class. Jack was inattentive when homework from the previous evening was reviewed. Closer investigation revealed that Jack ignored the review sessions because he was not doing the assignments. To increase the amount of homework completed, Mr. Vogl decided to use positive reinforcement. To evaluate the effectiveness of the intervention, he selected the AB design using the number of homework questions completed correctly as the dependent variable.

Over a baseline period of 5 days, Jack did 0 out of 10 (0/10) homework questions correctly each day. Because Jack frequently asked to listen to tapes in the French lab, Mr. Vogl decided to allow Jack to listen to tapes for 2 minutes for each correct homework question. Data collected during the intervention phase indicated an increase in the number of questions Jack answered correctly. Data analysis suggested that the intervention technique was effective.

Many teachers use AB designs to evaluate their students' progress.

Advantages and Disadvantages

The primary advantage of an AB design is its simplicity. It provides the teacher with a quick, uncomplicated means of comparing students' behavior before and after implementation of some intervention or instructional procedure, making instruction more systematic.

The disadvantage of the AB design is that it cannot be used to make a confident assumption of a functional relation. Although the data may show an increase or decrease in the behavior during the intervention phase, thus indicating effectiveness of the intervention, this design does not provide for a replication of the procedure. Therefore, the AB design is vulnerable to confounding variables or coincidental events. This is illustrated in the following example.

Miss Harper Conducts Research

As part of her initial student teaching assignment, Miss Harper was required to carry out a simple research project using an AB design. She decided to use Ralph's staying in his seat as her dependent variable. (Remember Ralph from Chapter 1.) Miss Harper collected baseline data for several days and determined that Ralph stayed in his seat for periods varying from 20 to 25 minutes during the 1-hour reading class. She prepared to intervene, choosing as her independent variable points exchangeable for various activities that Ralph enjoyed. When Professor Grundy made a visit soon after intervention began, Miss Harper met him at the door in a state of high excitement.

"It's working, Professor!" gloated Miss Harper. "Look at my graph! Ralph was absent the first 2 days of this week, but since he's been back and I've been giving him points, he's been in his seat 100% every day. Do you think I'll get an A on my project?"

"It's working professor. He's staying in his seat!"

Professor Grundy inspected Miss Harper's graph and agreed that her procedure appeared to be effective. He then sat down in the rear of the classroom to observe. After a few minutes, during which Ralph indeed stayed in his seat, Professor Grundy attracted Miss Harper's attention and called her to the back of the room.

"Miss Harper," he asked gently, "did it not occur to you that the heavy cast on Ralph's leg might have some effect on the amount of time he spends in his seat?"

Reversal Design

The **reversal design** is used to analyze the effectiveness of a single independent variable. Commonly referred to as the **ABAB design,** this design involves the sequential application and withdrawal of an intervention to verify the intervention's effects on

ABAB is a research design. A functional relationship can be demonstrated.

a behavior. By repeatedly comparing baseline data to data collected during application of the intervention strategy, the researcher can determine whether a functional relation exists between the dependent and independent variables.

IMPLEMENTATION

The reversal design has four phases: A, B, A, and B:

- A (baseline 1): the initial baseline during which data are collected on the target behavior under conditions existing before the introduction of the intervention.
- B (intervention 1): the initial introduction of the intervention selected to alter the target behavior. Intervention continues until the criterion for the target behavior is reached or a trend in the desired direction of behavior change is noted.
- A (baseline 2): a return to original baseline conditions, accomplished by withdrawing or terminating the intervention.
- B (intervention 2): the reintroduction of the intervention procedure.

Data collected using a reversal design can be examined for a functional relation between the dependent and independent variables. Figure 5–6 demonstrates a functional relation between the dependent and independent variables, said to exist if the second set of baseline data returns to a level close to the mean in the original A phase or if a trend is evident in the second A phase in the opposite direction of the first B phase. Figure 5–7 does not demonstrate the existence of a functional relation.

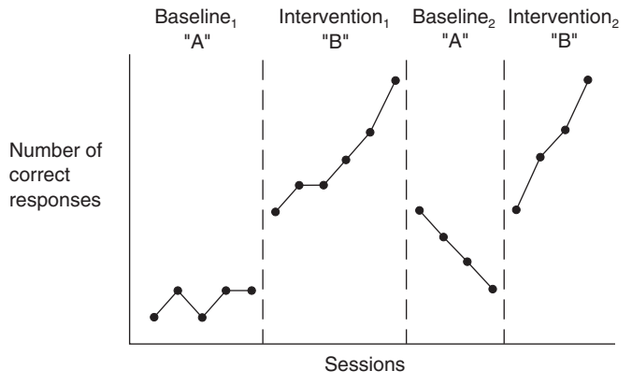


FIGURE 5–6 ***Reversal design graph that demonstrates a functional relationship between variables.***

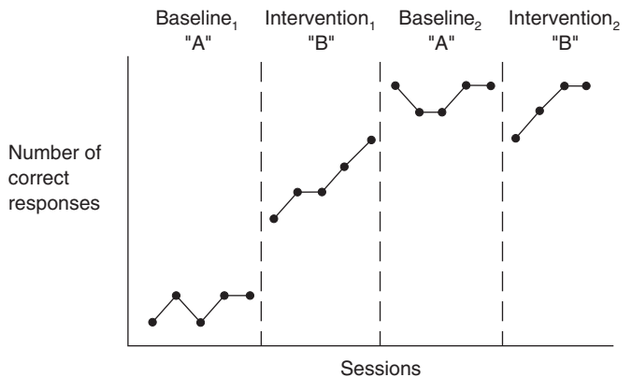


FIGURE 5–7 ***Reversal design graph that does not demonstrate a functional relationship between variables.***

Cooper (1981, p. 117) stated that researchers need three pieces of evidence before they can say that a functional relation is demonstrated: (1) *prediction*: the instructional statement that a particular independent variable will alter the dependent variable—for example, the contingent use of tokens to increase the number of math problems Michael completes; (2) *verification of prediction:* the increase (or decrease) in the dependent variable during the first intervention phase, and the approximate return to baseline levels of performance in the second A phase; and (3) *replication of effect:* the reintroduction of the independent variable during a second B phase resulting again in the same desired change in behavior.

The reversal design is a research design that allows the teacher to assume a functional relation between independent and dependent variables. The second baseline and intervention phases, with conditions identical to those of the first, provide an opportunity for replication of the effect of the intervention on the target behavior. It is unlikely that confounding variables would exist simultaneously with repeated application and withdrawal of the independent variable. The reversal design, however, is not always the most appropriate choice. The reversal design should not be used in the following cases:

1. When the target behavior is dangerous, such as aggressive behavior directed toward other students or self-injurious behavior. Because the reversal design calls for a second baseline condition to be implemented after a change in the target behavior rate, ethical considerations would prohibit withdrawing a successful intervention technique.
2. When the target behavior is not reversible. Many academic behaviors, for example, are not reversible, because the behavior change is associated with a learning process. Under such conditions, a return to baseline performance is not feasible. Knowledge that $4 \times 3 = 12$, for example, is not likely to be "unlearned." At least, we would like to think not.

It would be unethical to withdraw, for purposes of research, an intervention that stopped a student with severe disabilities from banging her head on the floor.

Graphic Display

The reversal design calls for four distinct phases of data collection. Figure 5–8 illustrates the basic reversal design. (Note that ABAB is derived from the labeling of each baseline period as an A phase and each intervention period as a B phase.)

Design Variations

Variations of the reversal design can be found in the literature. The first variation does not involve a change in the structure of the design, but simply shortens the length of

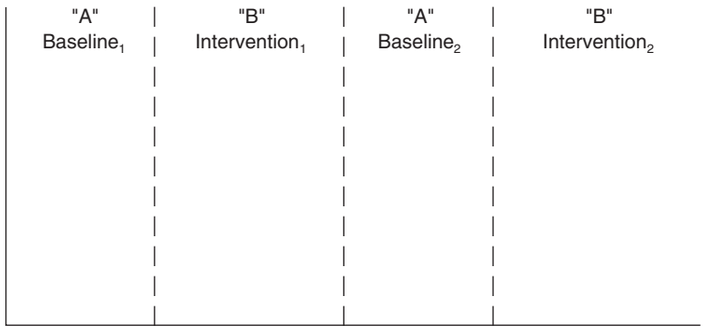


FIGURE 5–8
Basic reversal design format.

the initial baseline (A) period. This format of the design is appropriate when a lengthy baseline period is unethical, as when the behavior is dangerous, or not called for, in the case of a student who is not capable of performing the target behavior to any degree.

A second variation of the reversal design omits the initial baseline entirely. This BAB variation is considered if the target behavior is obviously not in the student's repertoire. When this design is used, a functional relation between the dependent and independent variables can be demonstrated only in the second intervention (B) phase.

Research Application

Researchers often use the ABAB design. Levendoski and Cartledge (2000) employed it to determine the effectiveness of a self-monitoring procedure for time on task and academic productivity with elementary-age students with emotional disturbance. The four boys were given self-monitoring cards at the beginning of each math period. They were told that each time they heard the bell (every 10 minutes), they should "Ask yourself . . . am I doing my work?" They were then to mark yes or no on their card.

Figure 5–9 shows the results of this intervention for time on task for one of the boys. During the baseline condition when the self-monitoring card was not being used, he was on task an average of 45%. Once the intervention was in place, his time on-task average increased to 93%. During the return to baseline phase, his on-task average returned to 34%, and then increased again to an average of 96% during reintroduction of the self-monitoring card. Examination of the graph clearly shows that when the student was using the self-monitoring card, his time on task increased. Note that phases one and two are replicated by phases three and four, allowing a determination of a functional relation.

Umbreit, Lane, and Dejud (2004) used an ABAB design to evaluate the effects of an intervention to increase on-task behavior of a general education fourth-grade student. During independent work assignments, Jason's off-task behavior included talking to other students, kicking his seat or the one in front of him, or wandering around the room. The teacher determined that this behavior occurred when he completed his assignment. Jason said he finished quickly because his assignments were "almost always too easy." During baseline phases, Jason received the same math and reading assignments as the rest of the

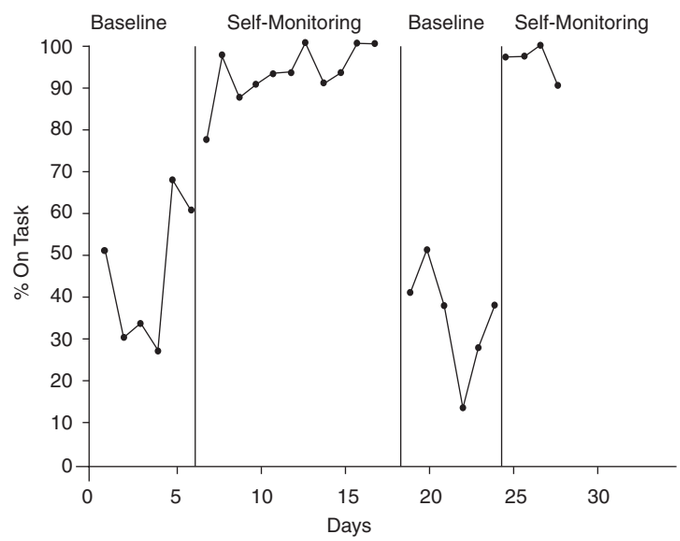


FIGURE 5–9
Research application of the reversal design.
Note: From "Self-Monitoring for Elementary School Children With Serious Emotional Disturbances: Classroom Applications for Increased Academic Responding," by L. Levendoski & G. Cartledge, 2000, *Behavioral Disorders.* Copyright 2000 by Council for Children with Behavioral Disorders. Reprinted by permission.

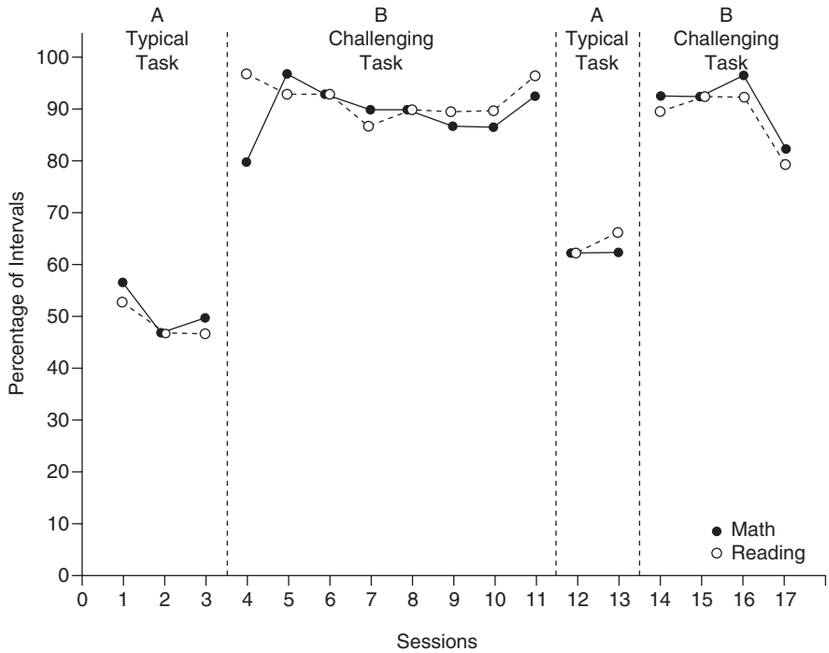


FIGURE 5–10
Research application of the reversal design.
Note: From "Improving Classroom Behavior by Modifying Task Difficulty: Effects of Increasing the Difficulty of Too-Easy Tasks," by J. Umbreit, K. Lane, & C. Dejund, 2004, *Journal of Positive Behavior Interventions, 6*, 13–20. Copyright 2004 by PRO-ED, Inc. Reprinted by permission.

class. During intervention phases, he received assignments that were more challenging (assignments approximately 2 weeks further along in the curriculum). Data for his on-task behavior were recorded using 30-second interval recording. As shown in Figure 5–10, during the first baseline condition (typical task), on-task behavior occurred approximately 50% of the intervals in both math and reading. During the first intervention phase, on-task behavior increased to an average of 89% in math and 92% in reading. During the second baseline phase, on-task behavior decreased to 63% during math and 65% during reading. In the final phase, on-task behavior increased to an average of 91% during math and reading. Comparison of the first baseline and first intervention phases shows an increase in on-task behavior when the challenging tasks were assigned. This effect was replicated during the second baseline and second intervention phases. This replication allows the assumption of a functional relation between the challenging tasks (independent variable) and on-task behavior during independent academic work (dependent variable).

Teaching Application

The following vignette illustrates the use of an ABAB design in the classroom.

Jill Learns Not to Suck Her Thumb

Ms. Kimball, a kindergarten teacher with 27 pupils, recently designed an effective intervention program to reduce her student Jill's thumb sucking. She decided that a reversal design would allow her to determine the existence of a functional relation between the change in behavior and the selected intervention procedure. Ms. Kimball chose a time sampling observation procedure. She looked at Jill every 10 minutes and marked a 1 on the data sheet if Jill was sucking her thumb and a 2 if she wasn't.

During the baseline condition (the first A phase), Ms. Kimball noted that Jill had her thumb in her mouth an average of 8 out of the 12 observations during a 2-hour period. Ms. Kimball decided to make a chart for Jill and to put a "smelly

sticker" on the chart whenever Jill was not sucking her thumb at the end of an interval. After the intervention was applied (the first B phase), thumb sucking occurred at the end of only three intervals, on average. To determine if a functional relation existed between the intervention and the behavior, Ms. Kimball returned to baseline (A) conditions. Jill no longer got smelly stickers, and the target behavior immediately returned to its previous level. Reinitiating the B condition immediately brought the behavior back to the lower level. Ms. Kimball felt confident the intervention had changed Jill's behavior.

Advantages and a Disadvantage

As the preceding applications show, the reversal design offers the advantages of simplicity and experimental control. It provides for precise analysis of the effects of a single independent variable on a single dependent variable.

The primary disadvantage of this design is the necessity for withdrawing an effective intervention in order to determine whether a functional relation exists. Even if the target behavior is neither dangerous nor irreversible, it often seems foolish to teachers to stop doing something that is apparently working.

Changing Criterion Design

The **changing criterion design** evaluates the effectiveness of an independent variable by demonstrating that a behavior can be incrementally increased or decreased toward a terminal performance goal. This design includes two major phases. The first phase (as in all single-subject research designs) is baseline. The second phase is intervention. The intervention phase is composed of subphases. Each subphase has an interim criterion toward the terminal goal. Each subphase requires a closer approximation of the terminal behavior or level of performance than the previous one. The student's performance thus moves incrementally from the baseline level to the terminal objective.

The changing criterion design is particularly appropriate when the terminal goal of behavior change is considerably distant from the student's baseline level. For example, if the goal is for the student to read 60 sight words, and her baseline level of performance is 5 words, it is probably unreasonable for the teacher to instruct and for her to learn all 55 words at once. It is better instructional and reinforcement practice for her to acquire a smaller number of words at a time. Similarly, if the goal is for the student to remain in his seat for 40 continuous minutes so he can be successful in an inclusive class, and his baseline level of performance is 5 continuous minutes, it is probably unreasonable to expect him to be able to master the entire 40 minutes at one time. It is more within his reach, and will provide many more opportunities for reinforcement, if he is brought gradually to the terminal goal of 40 continuous minutes in his seat.

The changing criterion design is well suited for measuring the effectiveness of a shaping procedure (see Chapter 9). This design is also useful when the teacher wants to accelerate or decelerate behaviors measured in terms of frequency, duration, latency, or force.

Implementation

The first step in implementing the changing criterion design is to gather baseline data in the same manner used in other single-subject designs. After a stable baseline has been established, the teacher must determine the level of performance change that will

be required for each subphase during intervention. The choice of the first interim level of performance may be determined using one of several techniques:

1. The interim criterion for performance can be set at, and then increased by, an amount equal to the mean of the stable portion of the baseline data. This technique is appropriate when the goal of the behavior-change program is to increase a level of performance and when the student's present level is quite low. For example, if a teacher wants to increase the number of questions a student answers and the student's mean baseline level of correct responses was two, that teacher might set two correct answers as a first interim criterion. Each subsequent subphase would then require two additional correct answers.
2. Interim criteria for performance can be set at half the mean of the baseline. If during the first intervention subphase raising the criterion by the mean of the baseline would make the task too difficult for the student, then raising it half that much may be appropriate. If the student's performance during the first intervention subphase is higher than a criterion equal to the mean of the baseline, the interim criterion may be raised by a level twice that of the mean of the baseline.
3. Interim criteria can be based on selecting the highest (or lowest, depending on the terminal objective) level of baseline performance. This is probably most appropriate for use with social behavior, such as out-of-seat or positive peer interactions, rather than for an academic behavior. The assumption is that if the student were able to perform at that high (or low) level once, the behavior can be strengthened (or weakened) and maintained at the new level.
4. Interim criteria can be based on a professional estimate of the student's ability. This procedure is particularly appropriate when the student's present level of performance is zero.

Regardless of the technique a teacher uses to establish the initial criterion, the data collected should be used to evaluate whether the amount of criterion change for each subphase is appropriate for a particular student.

The next step in implementing the changing criterion design is to begin the intervention phases. In each phase, if the student performs at least at the level of the interim criterion, the teacher provides reinforcement. It is important for the teacher to analyze the appropriateness of the selected interim level of performance during the initial intervention phase. If the student does not meet the criterion after a reasonable number of trials, the teacher should consider decreasing the interim level of performance required for reinforcement. Conversely, the teacher should consider adjusting the interim level of performance required for reinforcement if the student attains the goal too easily.

After the student has reached the established level of performance in a predetermined number of consecutive sessions (usually two, or in two out of three consecutive sessions of a subphase), the level of performance required for reinforcement should be adjusted in the direction of the desired level of performance for the overall behavior-change program. Each successive interim level of performance should be determined using the same mathematical difference established at the first interim level of performance. That is, the behavior-change program should reflect a uniform step-by-step increase or decrease in criterion level. This process is continued until

1. The behavior is increased to a 100% level or decreased to a 0% level of performance, or
2. The final goal established by the teacher in the behavioral objective is attained.

Changing criterion designs enable teachers and researchers to establish functional relationships.

A functional relation between the dependent and independent variable is demonstrated if the student's performance level matches the continually changing criterion for performance and reinforcement specified by the teacher (Kazdin, 1998; Richards, Taylor, Ramasamy, & Richards, 1999). This method of assessing a functional relation is based on the view that repeated matching to a changing criterion represents instances of replication. Each subphase with its interim criterion serves as the baseline for the increased (or decreased) criterion of the next subphase (Cooper, Heron, & Heward, 2007; Hartmann & Hall, 1976). Generally, a student must meet the established criteria in at least three consecutive phases before the assumption of a functional relation is valid.

Graphic Display

The basic changing criterion design format is similar to the one used for the AB design. A baseline phase is followed by the intervention phase, with a dashed vertical line separating the two conditions and each subphase. Figure 5–11 shows that the data for the intervention phase are identified according to the level of performance selected for reinforcement. The procedure for graphing the data calls for connecting data points within each subphase. Data points collected in different interim phases or subphases are never connected. The magnitude of student behavior necessary for consequation (delivery of reinforcement) should be clearly identified at each level of the intervention phase (see Figure 5–11).

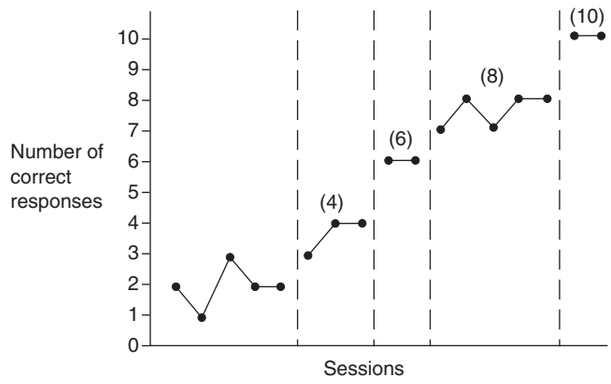
Research Application

Hall and Fox (1977) used the changing criterion design to increase the number of math problems correctly solved by a child with a behavior disorder. Under baseline conditions, the student demonstrated a mean level of performance of one math problem.

The first interim level of performance was established at the next whole number greater than the mean baseline performance (2). If the student met this level of performance, he was allowed to play basketball. If the student failed to reach the criterion, he had to stay in the math session until the problems were solved correctly. Figure 5–12 shows that this process was continued until 10 math problems were solved correctly.

Ellis, Cress, and Spellman (1992) used a changing criterion design to demonstrate an increase in self-managed independent exercise by students with moderate and severe mental retardation. Three of the students were taught to use a treadmill. The student

FIGURE 5–11
Basic changing criterion design format.



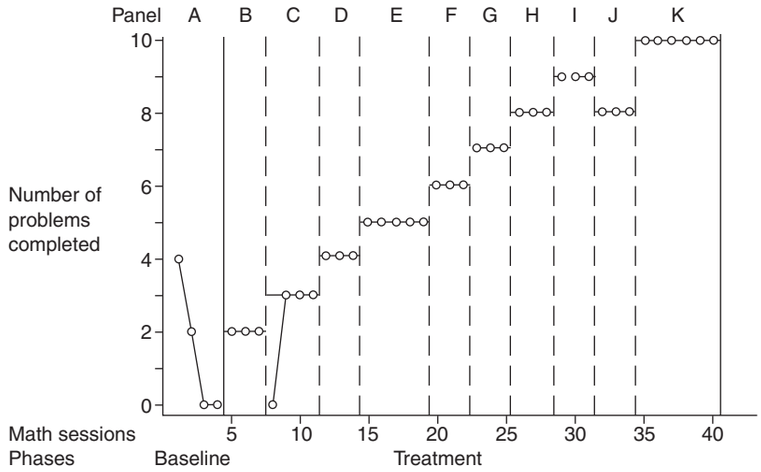


FIGURE 5–12
Research application of the changing criterion design.
Note: From "Changing Criterion Designs: An Applied Behavior Analysis Procedure," by R. V. Hall & R. G. Fox, 1977, in B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New Developments in Behavioral Research: Theory Method and Application.* Copyright 1977 by Lawrence Erlbaum Associates, Inc., Publishers. Reprinted with permission of the authors and the publishers.

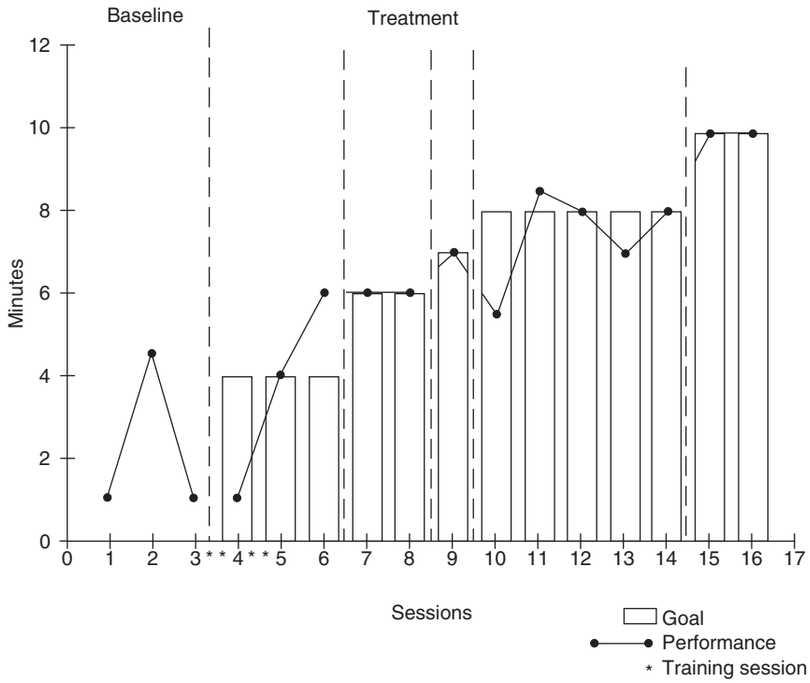


FIGURE 5–13
Application of the changing criterion design.
Note: From "Using Timers and Lap Counters to Promote Self-Management of Independent Exercise in Adolescents with Mental Retardation," by D. Ellis, P. Cress, & C. Spellman, 1992, *Education and Training in Mental Retardation.* Copyright 1992 by the Council for Exceptional Children. Reprinted by permission.

whose graph is presented in Figure 5–13 was a 16-year-old girl. Following a baseline taken during unsupervised sessions held twice a week for 2 weeks, durations were determined for the intervention subphases. Initial durations were near or slightly longer than the longest baseline performance. Interim criteria were then increased generally by 2 minutes once the previous goal was achieved in one or more sessions. (The interim criterion was raised by only 1 minute during one subphase to further demonstrate control over the behavior by the independent variable.) Students were taught to set a digital kitchen timer, and stickers served as token reinforcers for meeting interim criteria.

Students used the tokens to purchase exercise-related items (e.g., T-shirts, sweatbands, shorts) on a 5:1 ratio. The students displayed a systematic increase in the number of minutes of self-managed exercise from one subphase and interim criterion to the next. The authors attributed those sessions in which the student whose graph is presented did not reach her interim criterion to student errors in timer activation or lags between timer activation and exercise.

Certain procedural elements may increase the research credibility of the changing criterion design by enhancing experimental control:

1. Continuing with a subphase until a stable rate has been established
   For classroom use, maintaining a behavior at the interim criterion for two sessions (or two out of three sessions) before moving to the next subphase demonstrates sufficient control. Because, for research purposes, each subphase is seen as the baseline for the following subphase, the subphase may be continued until a stable rate has been established before starting the next subphase (Richards et al., 1999).
2. Altering the number of sessions in some subphases
   In Figure 5–12, three sessions at each interim criterion were generally maintained; however, this number of sessions was changed in some subphases. The lengths of subphases may vary with the behavior's remaining at criterion level as long as the criterion is in effect (Cooper et al., 2007; Richards et al., 1999). "It is stability after change has been achieved, and before introduction of the next change in criterion, that is crucial to producing a convincing demonstration of control" (Hartmann & Hall, 1976, p. 531).
3. Varying the increase (or decrease) in performance required in subphases
   In Figure 5–13, the third subphase criterion was set at an increase of 1 minute rather than 2 minutes. Varying the size of criterion changes provides a more convincing demonstration of experimental control (Cooper et al., 2007).
4. Requiring a change in a direction opposite to the terminal goal in one or more phases.

In Figure 5–12, in subphase J, a change in the criterion for reinforcement was made in the direction opposite the terminal objective. Returning the student's performance level to a previously mastered criterion demonstrates a reversal effect similar to that of a return to baseline condition in an ABAB design.

Teaching Application

Here is a teaching application of a changing criterion design to compare with the research application.

Claudia Learns to Sort by Color

Claudia was a student in Mr. Carroll's intermediate class for students with moderate mental retardation. Mr. Carroll was trying to teach Claudia to sort objects rapidly by color. Claudia could perform the task, but she did it too slowly. Mr. Carroll decided to use a changing criterion design to evaluate the effectiveness of a positive reinforcement procedure. He established that Claudia's average baseline rate of sorting was 4 objects per minute. He set 6 per minute as the first interim criterion and 30 per minute as the terminal goal. Claudia earned a poker chip exchangeable for a minute's free time when she met the criterion. When Claudia met the criterion on two consecutive trials or opportunities, Mr. Carroll raised the criterion required for reinforcement by two. He continued to do this until Claudia sorted 30 objects per minute in order to earn her poker chip. Mr. Carroll concluded that there was a functional relation between the dependent and independent variables, because Claudia's behavior changed quickly each time the criterion was changed but did not change until then.

Advantage and Disadvantage

The advantage of the changing criterion design is that it can establish a functional relation while continually changing the behavior in a positive direction. There is no need to withdraw a successful intervention. Using the changing criterion design, however, necessitates very gradual behavior change. It may therefore be inappropriate for behaviors that require or lend themselves to rapid modification.

Multiple Baseline Design

When to use a multiple baseline design.

As indicated by its name, the **multiple baseline design** permits simultaneous analysis of more than one dependent variable. A teacher may experimentally test the effects of intervention (the independent variable) on

1. two or more behaviors associated with one student in a single setting, such as John's out-of-seat and talking-out behaviors in social studies class (*multiple baseline across behaviors*).
2. two or more students exhibiting the same behavior in a single setting, as in the spelling accuracy of Sara and Janet in English class (*multiple baseline across individuals*).
3. two or more settings in which one student is exhibiting the same behavior, such as Kurt's cursing during recess and in the school cafeteria (*multiple baseline across settings*).

The multiple baseline is the design of choice when the teacher is interested in applying an intervention procedure to more than one individual, setting, or behavior. The multiple baseline design does not include a reversal phase; therefore, it may be used when the reversal design is not appropriate: when the target behavior includes aggressive actions or when academic learning is involved.

"Baseline, directionality, stability, control—is this teaching or tennis?"

Implementation

A teacher using the multiple baseline design collects data on each dependent variable simultaneously. The teacher collects data under baseline conditions for each student, on each behavior, or in each setting. In establishing the data collection system, the teacher should select an ordinate scale that is appropriate for each of the variables involved in the program. To make data analysis possible, the same scale of measurement (for example, number of math problems completed correctly or percent of on-task behavior) should be used for each dependent variable.

After a stable baseline has been achieved on the first variable, intervention with that variable can be started. During the intervention period, baseline data collection continues for the remaining variables. Intervention on the second variable should begin when the first variable has reached the criterion established in the behavioral objective or when the data for the first variable show a trend in the desired direction as indicated by three consecutive data points. The intervention condition should be continued for the first variable, and baseline data should still be collected for any additional variables. This sequence is continued until the intervention has been applied to all the variables identified for the behavior-change program.

The data collected in a multiple baseline design can be examined for a functional relation between the independent variable and each of the dependent variables. The introduction of the intervention with the second and subsequent dependent variables constitutes a replication of effect. For example, after taking baseline data on Matt's on-task behavior in the special education resource class and in environmental science class, the teacher begins intervention in the resource class. Matt is presented with the contingency that if he is on task 85% of the times the teacher looks over at him, he will be able to reduce his homework assignment by 20%. The contingency goes into effect on Tuesday and continues for 4 days until his behavior meets this criterion. During these same 4 days the teacher has continued to take baseline data in the science class. Once Matt has reached the criterion in the resource class, the contingency is put into effect in the science class and continues to be in effect in the resource class. If Matt's on-task behavior is increased in the resource room and then increased in the science class, the teacher can say there is a functional relation between Matt's on-task behavior and earning a reduction in homework. There is a functional relation because the effect was first seen in the resource class and then replicated across settings in the environmental science class. A functional relation is assumed if each dependent variable in succession shows a change when, and only when, the independent variable is introduced.

Adjacent graphs should be examined to be sure that each successive intervention has an independent treatment effect on the appropriate dependent variable. Only the first independent variable should be affected by the first intervention. A change in the second and succeeding dependent variables should be seen only when the intervention is applied to them as well. Figure 5–14 shows an example of a functional relation, whereas Figure 5–15 does not. In Figure 5–15, the second dependent variable begins an upward trend when the intervention is introduced for the first variable, showing that the relations between variables are not discrete, or independent.

Graphic Display

When using the multiple baseline design, the teacher should plot the data collected using a separate axis for each of the dependent variables to which intervention was

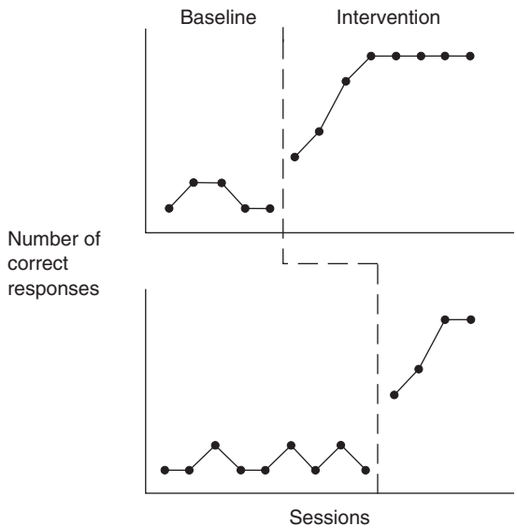


FIGURE 5–14
Data from a multiple baseline design that reflect a functional relationship.

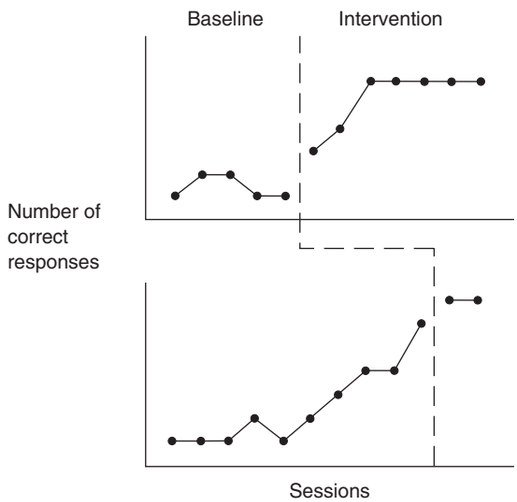


FIGURE 5–15
Data from a multiple baseline design that do not reflect a functional relationship.

applied (individuals, behaviors, or situations). Figure 5–16 shows a composite graph of a multiple baseline design.

Research Applications

Across Behaviors

Higgins, Williams, and McLaughlin (2001) used a multiple baseline design across behaviors to determine if a token reinforcement program could decrease the high rates of three inappropriate behaviors of an elementary student with learning disabilities. The three behaviors were out of seat, talking out, and poor seat posture. The student earned tokens for not exhibiting the three behaviors during daily independent work sessions. The contingency was first in effect for his talking-out behavior, then placed on his

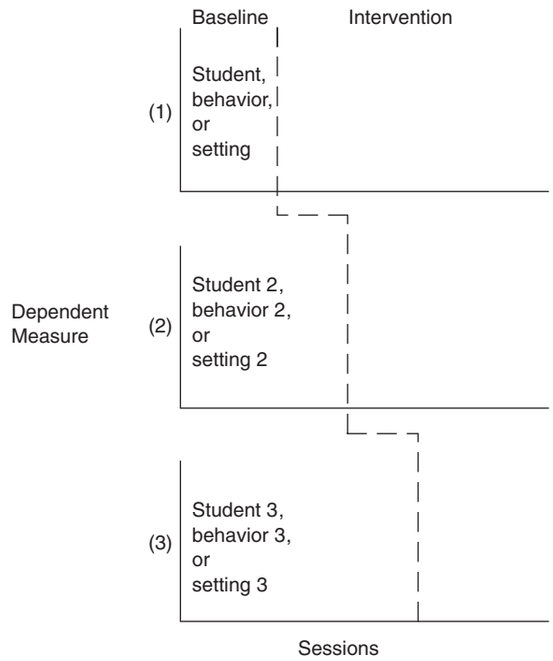


FIGURE 5–16
Basic multiple baseline design format.

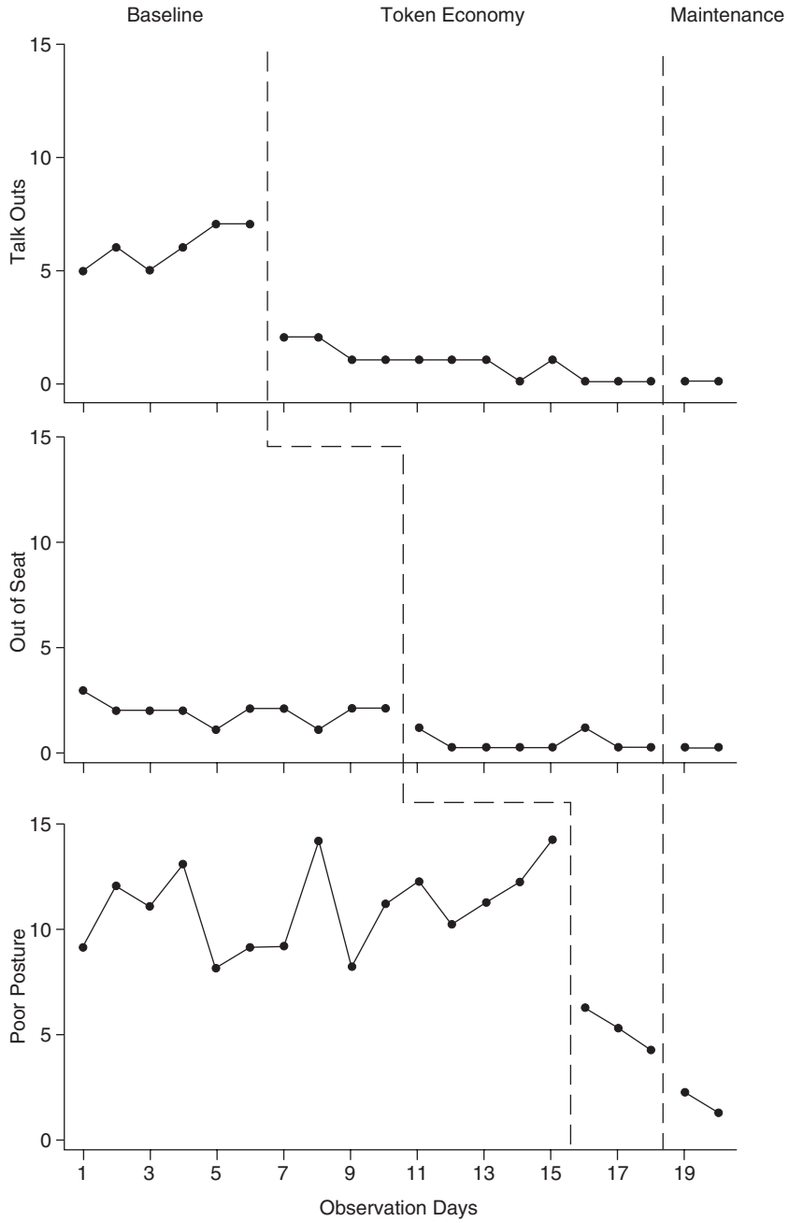
out-of-seat behavior, and finally on his seat posture. Figure 5–17 indicates that the reinforcement program was effective for reducing each behavior in succession. A functional relation between each dependent variable (the behaviors) and the independent variable (the token economy) can be said to exist because the successful use of the token economy was replicated across the three behaviors.

Across Individuals

Buggey (2005) used a multiple baseline design across individuals to evaluate the effectiveness of videotaped self-modeling to instruct positive behaviors (social interactions) and reduce tantrums with middle school boys diagnosed with Asperger syndrome. Social initiations were defined as unsolicited verbalizations addressed to peers or staff. A role-playing script was written and peers from the school were asked to create a movie. Scenes showed Roy or Tommy walking up to students and asking social questions followed by a brief discussion of favorite activities. A 3-minute video was created and shown to the students in the classroom prior to the start of classes. Data were collected during lunch, recess, and free time, each of which was 30 minutes. As shown in Figure 5–18, the results indicated both boys made immediate gains in the frequency of their social initiations and those gains were maintained. Roy, from no social initiations during baseline, rose to 4.0 initiations per day and maintained at an average of 4.4 a day. Tommy had two social imitations in the 12 days of baseline for a daily mean of .17. This increased to 3.8 and maintained at 4.25 initiations per day. Since the success of videotaped self-modeling with Roy was replicated with Tommy, one can say there was a functional relation between the dependent and independent variables.

Researchers have used the multiple baseline design to look at the behavior of more than one student at a time. They identify an individual unit or single case made up of more than one person as the unit of analysis. They have identified dyads in which pairs of students were identified as a single case (e.g., Duker, Hensgens, & Venderbosch, 1995;

FIGURE 5–17 ***Graph of multiple baseline design across behaviors.***



Note: From "The Effects of a Token Economy Employing Instructional Consequences for a Third-Grade Student with Learning Disabilities," by J. Higgins, R. Williams, & T. McLaughlin, 2001, *Education and Treatment of Children.* Copyright 2001 by Family Services of Western PA. Reprinted by permission.

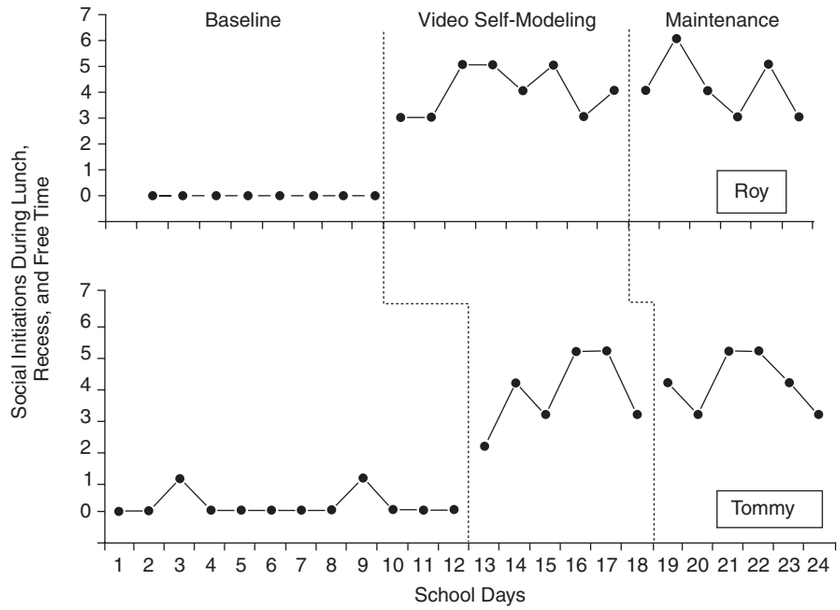


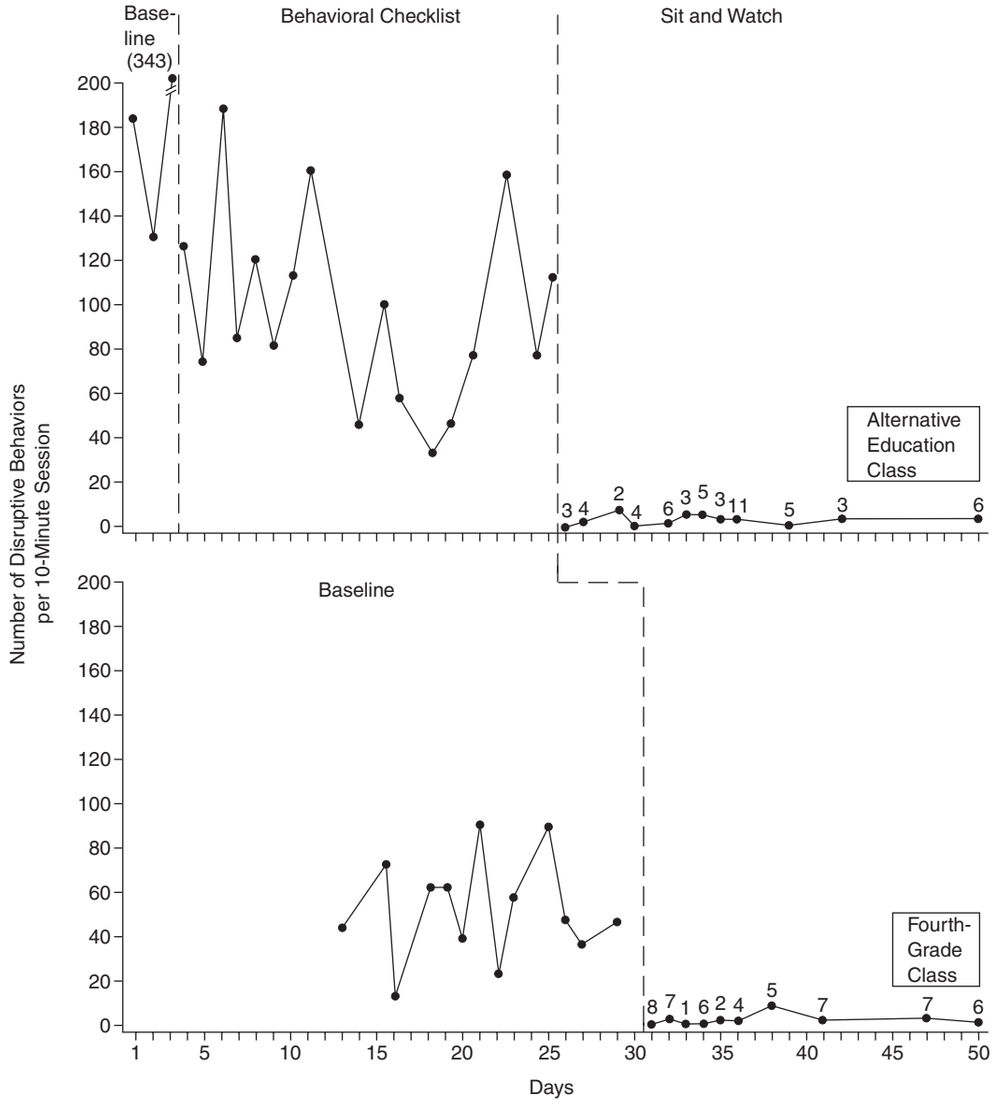
FIGURE 5–18
Graph of a multiple baseline design across individuals.
Note: From "Video Self-Modeling Applications with Students with Autism Spectrum Disorder in a Small Private School Setting," by T. Buggey, 2005, *Focus on Autism and Other Developmental Disabilities, 20*, 52–63. Copyright 2005, by the Hammill Institute on Disability, Inc. Reprinted with permission.

Loncola & Craig-Unkefer, 2005), pairs of students and paraeducators (e.g., Malmgren, Causton-Theoharis, & Trezek, 2005; Schlosser, Walker, & Sigafoos, 2006), and pairs of students and parents (e.g., Singh, Lancioni, Winton, et al., 2006). Researchers have identified even larger units such as small instructional groups of students as a single case (e.g., Fueyo & Bushell, 1998; Hawken, MacLeod, & Rawlings, 2007), a whole class as an individual unit (e.g., Kohler, Strain, Hoyson, & Jamieson, 1997; White & Bailey, 1990), and entire schools as a unit (e.g., Scott, 2001; Van Houten, Van Houten, & Malenfant, 2007). The reporting and graphing of student performance is usually reported as either an average of the target behavior as performed by group members as a whole, or the performance of individual members within the group. White and Bailey (1990) assessed the effectiveness of a "sit and watch" procedure, a form of time-out, on the disruptive behavior (noncompliance, aggression, and throwing objects) of 30 regular fourth-grade students and 14 students in an alternative education class for fourth- and fifth-grade boys with severe behavior problems. Students who engaged in disruptive behavior were required to sit and watch the other students playing for 3 minutes. During each observation interval, the observers recorded each instance of disruptive behavior on hand counters. Figure 5–19 shows data for both classes. The numbers above the data points represent the number of times "sit and watch" was implemented. A functional relation is assumed because the disruptive behavior decreased for each group of students when the intervention was in place.

Across Settings

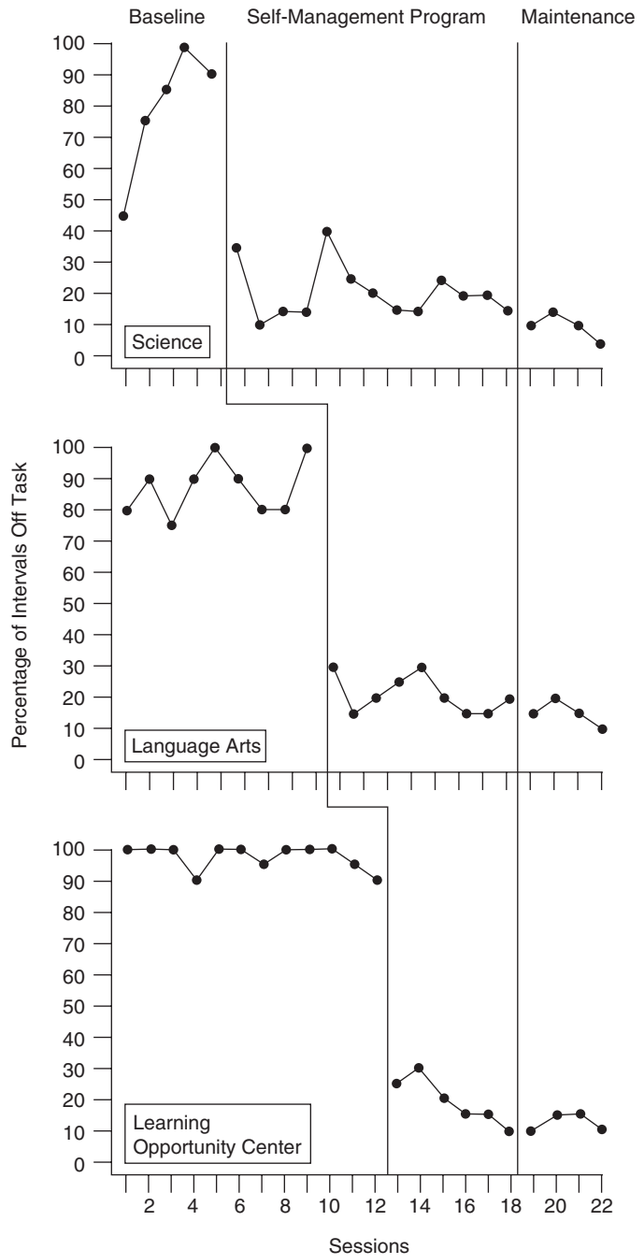
Dalton, Martella, and Marchand-Martella (1999) used a multiple baseline design across settings to evaluate the effects of a self-management program on the off-task behavior of two eighth-grade boys with learning disabilities. Off task was operationally defined as (a) not in seat (buttocks were not on the seat of chair, feet did not have to be on the floor), (b) talking with others (student talking, whispering, or mouthing to others without permission), (c) interrupting others (passing a note, touching another student's body or possessions), (d) not working on assigned task (scribbling or doodling instead of

FIGURE 5–19 ***Use of a multiple baseline design across individuals in which an entire class was considered "an individual."***



writing, reading a magazine instead of the text), and (e) engaging in bodily movements unrelated to or interfering with assigned task (playing with pencil or ripping paper). Figure 5–20 presents Peter's graphed data. During baseline, "normal classroom procedures" were in place. These consisted of redirection, reprimand, removal from class, or detention. Peter's off-task behavior averaged 79% in science class, 87% in language arts, and 97% in learning opportunity center (study hall). The self-management program was first introduced in science class where time off task was reduced to an average of 17%, then in language arts where it was reduced to an average of 21%, and finally in study

FIGURE 5–20 ***Graph of a multiple baseline design across settings.***



hall where his off-task behavior decreased to an average of 16%. Notice that the self-management program was initially used in science class, then replicated in language arts, and replicated a second time in learning opportunity center. These successful replications allow for a conclusion that there was a functional relation between the dependent variable and the independent variable.

Teaching Application

This vignette illustrates the use of a multiple baseline design in the classroom.

Students Learn to Come to Class on Time

Ms. Raphael was a middle school English teacher. The students in all three of her morning classes consistently came late. She began to record baseline data on the three classes. She recorded the number of students in their seats when the bell rang. She found that an average of five students in the first class, four in the second, and seven in the third class were in their seats. Ms. Raphael then began recording an extra-credit point in her grade book for each student in the first class who was in his or her seat when the bell rang. Within a week, 25 students were on time and in their seats. The baseline data for the other classes showed no change during this first intervention. When she began giving extra-credit points in the second class, the number of students on time increased immediately and dramatically. After a week, she applied the intervention in the third class with similar results. Ms. Raphael had accomplished two things: she had succeeded in getting her classes to arrive on time and she had established a functional relation between her intervention (the independent variable) and her students' behavior (the dependent variable).

Advantages and Disadvantages

A problem with the multiple baseline design—and a suggested solution.

The multiple baseline design can establish a functional relation without withdrawing the intervention, as is necessary in a reversal design, and without gradual alteration, as is required in a changing criterion design. These advantages make it a particularly useful design for classroom use. The multiple baseline design does, however, have some limitations. This design requires that the researcher apply the intervention to several students, behaviors, or settings, which may not always be practical. The multiple baseline design also requires collecting baseline data over extended periods, particularly baseline data for the second and subsequent dependent variables. When the student cannot perform the behavior at all or access to additional settings is not available or practical, collecting daily baseline data may take more time than is actually warranted or may not be possible. The multiple probe technique has been suggested as a reasonable solution to this situation (Horner & Baer, 1978; Poling, Methot, & LeSage, 1994). In this variation of the multiple baseline design, data are not continuously collected on the behaviors (or students or settings) on which intervention is not being conducted. Rather, probe trials (single trials under baseline conditions) or a probe session (more than one trial under baseline conditions) are conducted intermittently on these subsequent behaviors to verify that the student still cannot perform the behavior or to record any changes in his ability before intervention. While using the intervention with behavior 1 (or with student 1 or in setting 1), the teacher intermittently probes behaviors 2 and 3. When behavior 1 reaches the criterion, one or more probe sessions are conducted on all three behaviors. Then intervention is begun on behavior 2. Postcheck probes are conducted on behavior 1 to establish that the change in behavior is being maintained, and baseline probes continue on behavior 3. When behavior 2 reaches the criterion, one or more probe sessions are conducted on all three behaviors. Then intervention is begun on behavior 3, while postcheck probes are conducted on behaviors 1 and 2.

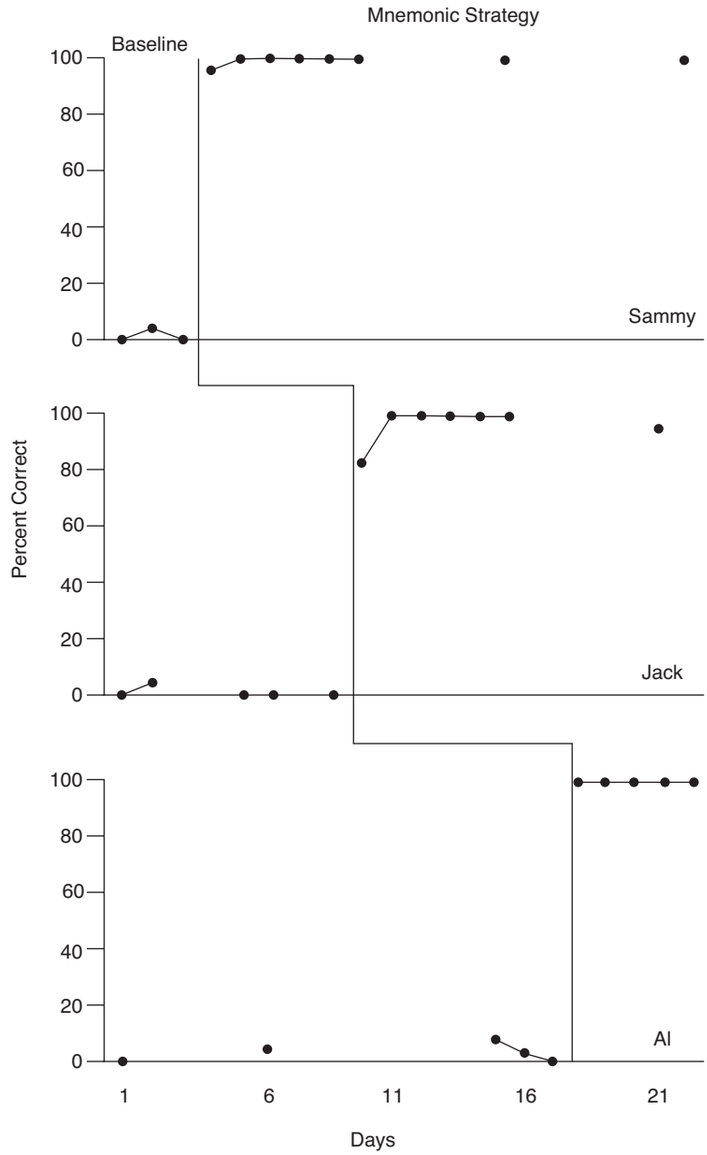


FIGURE 5–21
Use of a multiple probe design.
Note: From "Teaching Students with Severe Emotional or Behavioral Disorders to Use a Musical Mnemonic Technique to Solve Basic Division Calculations," by T. Cade & P. Gunter, 2002, *Behavioral Disorders*. Copyright 2002 by the Council for Children with Behavioral Disorders. Reprinted by permission.

Cade and Gunter's (2002) use of a multiple probe is presented in Figure 5–21. In this study three students, ages 12 to 14, diagnosed with emotional or behavioral disorders were taught to use a mnemonic to solve basic division calculations. Permanent product data were collected on a worksheet of 24 division-by-7 facts. Initially, baseline data were collected on all three students. Sammy quickly established a stable trend (after 3 consecutive days), and baseline sessions for the other two boys were continued as probes. For Jack, baseline probes were taken on days 1, 2, 5, and 6. For Al, baseline probes were taken on days 1, 6, 15, 16, and 17. Use of the mnemonic was begun with each successive student as the previous student reached a criterion of five sessions at 100% correct. The use of baseline probes eliminated the necessity of recording a

continuous baseline for Jack and Al on a behavior it was known they could not perform, but allowed for monitoring any potential change in their behavior before intervention.

The multiple baseline design is inappropriate in two specific situations:

1. when the target behavior calls for immediate action. The multiple baseline design calls for a considerable delay in delivery of the intervention procedure for the second and subsequent dependent variables.
2. when the behaviors selected for intervention are not independent. In such a case, intervention with one behavior will bring about a change in the related behavior; therefore, the teacher will be unable to evaluate clearly the effects of the procedure. For example, if two behaviors targeted for a student are cursing and fighting, the teacher might find that after the student's cursing decreases, fewer fights occur. In this case, the two behaviors are clearly not independent.

Alternating Treatments Design

In contrast to the multiple baseline design, which uses a single independent variable and multiple dependent variables, the **alternating treatments design** (Barlow & Hayes, 1979; Richards et al., 1999) allows comparison of the effectiveness of more than one treatment or intervention strategy on a single dependent variable. For example, using this design, the teacher can compare the effects of two reading programs on a student's reading comprehension ability or the effects of two behavior-reduction procedures on a student's talking out. The teacher can also examine the efficiency of three different types of symbols on a student's communication board. A number of different terms have been used to describe this design: multiple schedule design (Hersen & Barlow, 1976), alternating conditions design (Ulman & Sulzer-Azaroff, 1975), and multi-element baseline design (Sidman, 1960).

Implementation

The first step in setting up an alternating treatments design is to select the target behavior and two or more potential treatments. If the target behavior is social (for example, asking appropriate questions or remaining on task), it should be operationally defined. If the target behavior is academic, two or more representative samples of the behavior (for example, two or more equally difficult sets of division problems) should be selected, each designated for one of the intervention or treatment strategies.

As the name of this design implies, the treatments are implemented alternately or in rotation. The presentation of the treatments may be in random order, such as ABBABAAB (Barlow & Hersen, 1984). When two treatments are used, the student should be exposed to each treatment an equal number of times. If there are three treatments, a block rotation may be used. Each block consists of one presentation of each treatment; for example, ABC, BCA, CAB, ACB, BAC, CBA. If data are collected long enough, each possible order of presentation should be used at least once.

The alternating treatments can be used sequentially within a single session (A followed by B), or from one session to the next (A in the morning, B in the afternoon of the same day), or on successive days (A on Monday, B on Tuesday). The scheduling should be counterbalanced; that is, the treatment that was employed first in one session should be used second in the next session, the treatment employed in the morning on the first day should be used in the afternoon on the second day, and the treatment that was used on Monday the first week should be used on Tuesday of the second week.

(In research situations, similar counterbalancing is used to minimize the effects of other potential confounding variables such as the person administering the treatment and the location of the treatment.) This counterbalancing should control for the possibility of carryover and sequencing effects (Barlow & Hayes, 1979). In other words, by presenting the treatments in random order, the possible effects each treatment may have on the others will be minimized.

A distinctive discriminative stimulus, signal, or cue immediately preceding each treatment will make it clear to the student which condition is in effect. For example, the teacher might say, "This is treatment A" and "This is treatment B," or "Now we are going to use a number line" and "Now we are going to use counting chips." The teacher might also color code worksheets to indicate that a particular condition is in effect.

Graphic Display

The basic form of graphing the alternating treatments design is shown in Figure 5–22. As in all designs, baseline data are plotted first and separated from intervention data by a vertical broken line. The graph for the alternating treatments design differs from others in that several curves may be shown on each graph. The points for each treatment are connected only to other points for that treatment so that the data for each are displayed as separate lines, or curves.

If the data curve of one treatment is vertically separated from the other curves, it is said to be *fractionated*. This fractionation indicates that the treatments are differentially effective (Ulman & Sulzer-Azaroff, 1975).

The top graph in Figure 5–22 shows data that demonstrate an effective treatment. Treatment A is the more effective of the two treatments. The data curves are separated; they do not cross at any point other than at the very beginning of the intervention phase. The two curves are fractionated. Figure 5–22 also shows data that are not significantly different from one another. The middle graph shows two treatments, neither of which demonstrates control over the dependent variable; thus, neither is effective. The bottom graph shows two treatments that both demonstrate control over the dependent variable; thus, either is equally effective.

By visual inspection of the graphs, we may infer experimental control between one or more of the independent variables and the dependent variable.

> Because confounding factors such as time of administration have been neutralized (presumably) by counterbalancing, and because the two treatments are readily discriminable by subjects through instructions or other discriminative stimuli, differences in the individual plots of behavior change corresponding with each treatment should be attributable to the treatment itself, allowing a direct comparison between two (or more) treatments. (Barlow & Hayes, 1979, p. 200)

As described thus far, the alternating treatments design does not include a replication phase. Therefore, the case for the existence of a functional relation is relatively weak. To make a stronger case, a third phase can be instituted. In this phase the more effective of the treatments is applied to the behavior (or behavior sample) that was treated with the ineffective treatment during the intervention phase. If the second behavior then improves, a replication of the treatment has been accomplished and a functional relation demonstrated. Figure 5–23 shows such a three-phase variation of the design.

Research Application

Singh (1990) used an alternating treatments design to measure the comparative effectiveness of two error-correction procedures in reducing oral reading errors of students

FIGURE 5–22 ***Graphs of data collected using an alternating treatments design. Top graph indicates treatment A is the more effective treatment. Middle and bottom graphs indicate no difference between treatments.***

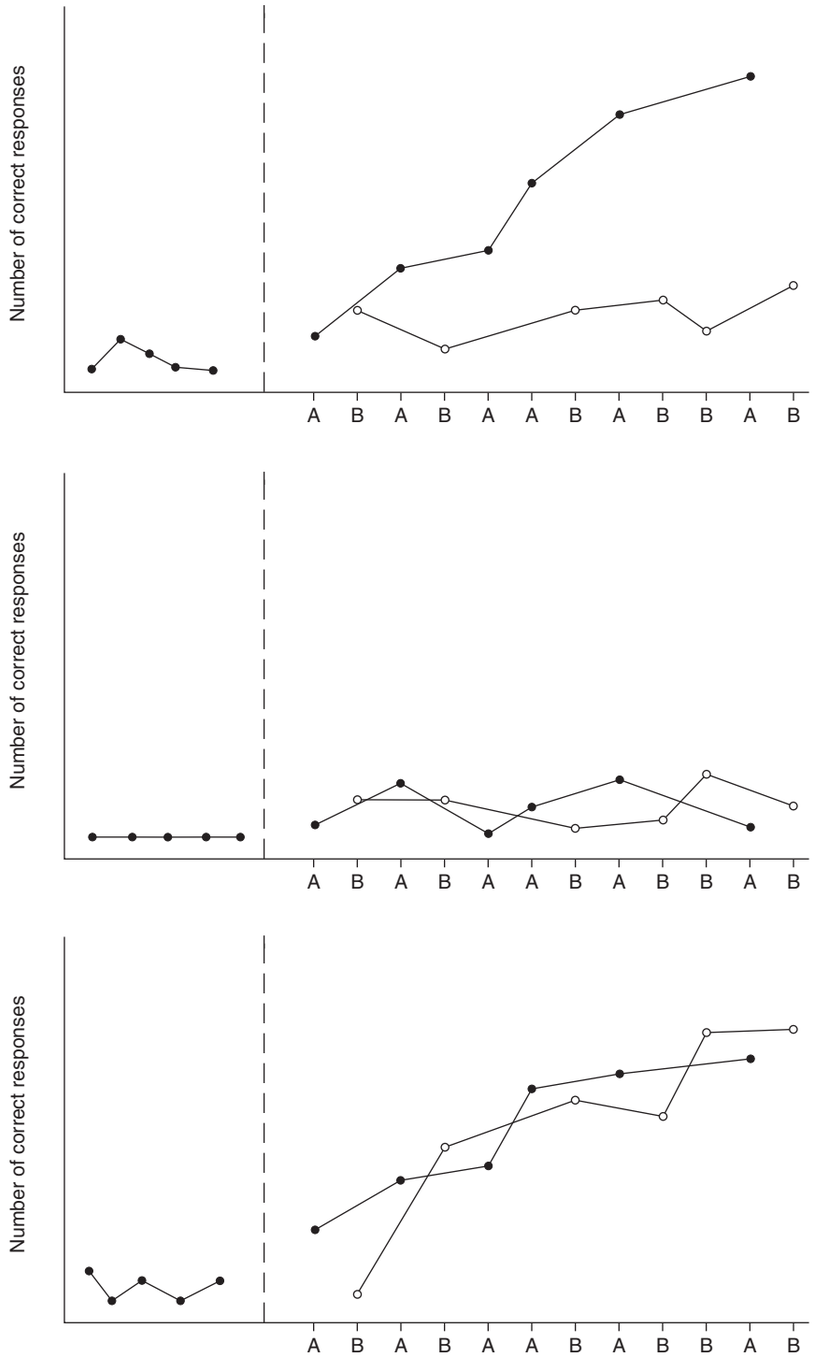
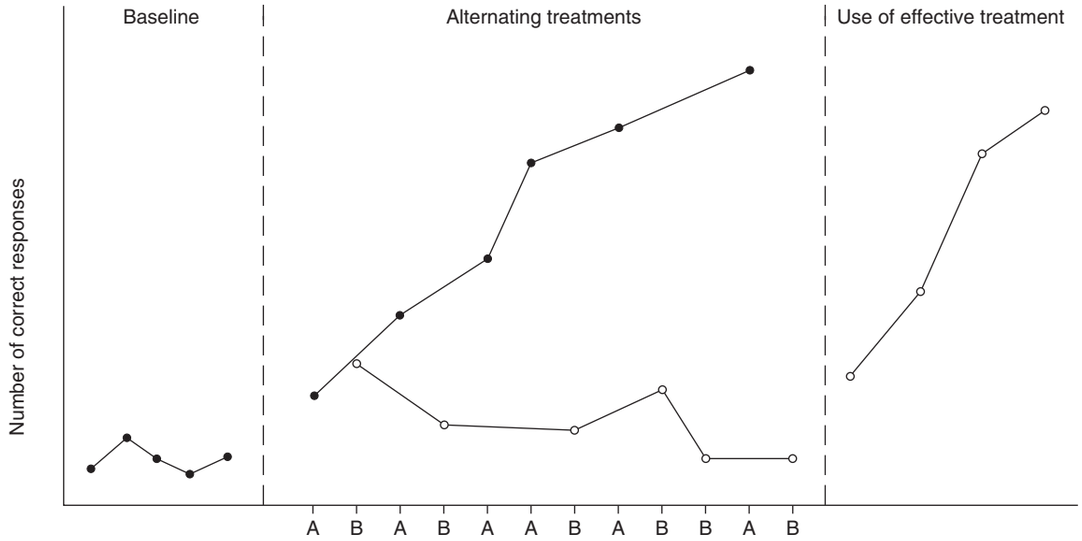


FIGURE 5–23 ***Three-phase alternating treatments design demonstrating functional relationship.***



with moderate mental retardation. Students read an unfamiliar 100-word passage orally three times each day. The correction procedures were word supply (the teacher supplied the correct word, the student repeated it once and continued to read) and sentence repeat (the student repeated the correct word after the teacher, read the rest of the sentence, then reread the entire sentence). As shown in Figure 5–24, the sentence-repeat procedure was more effective than the word-supply procedure (or a control—no correction) for each of the three students. In the third phase of the design the more effective procedure was used in all three daily sessions. This provided the replications needed to assess the presence of a functional relation.

Cihak, Alberto, Taber-Doughty, and Gama (2006) used an alternating treatments design to compare the effectiveness of static pictures (photographs) as prompts with video prompting. Within a group instruction format the middle school students with moderate mental retardation were taught using static picture prompts to use a bank card to make a purchase and using video prompts to withdraw money from an ATM (the prompting system and tasks were counterbalanced with three other students). During community-based instruction students performed one session (i.e., two trials) of each task sequence. A least-to-most prompt hierarchy was used until the student performed correctly without assistance. Figure 5–25 shows the data for percent of steps of the task analysis performed correctly by each group member. The data show that for each student, performance was low and stable during baseline and increased during intervention. During intervention the data paths for static pictures and for video prompts continually overlap, indicating that there was no difference in their effectiveness with any of the students. A probe 2 weeks later indicated that these results were maintained. (Note that this is not also a multiple baseline design because the intervention was begun with all group members at the same time.)

Using an alternating treatments design can help teachers individualize instruction.

Teaching Application

For teachers, the alternating treatments design can provide rapid and accurate feedback about the comparative effectiveness of various teaching techniques, as the following example shows.

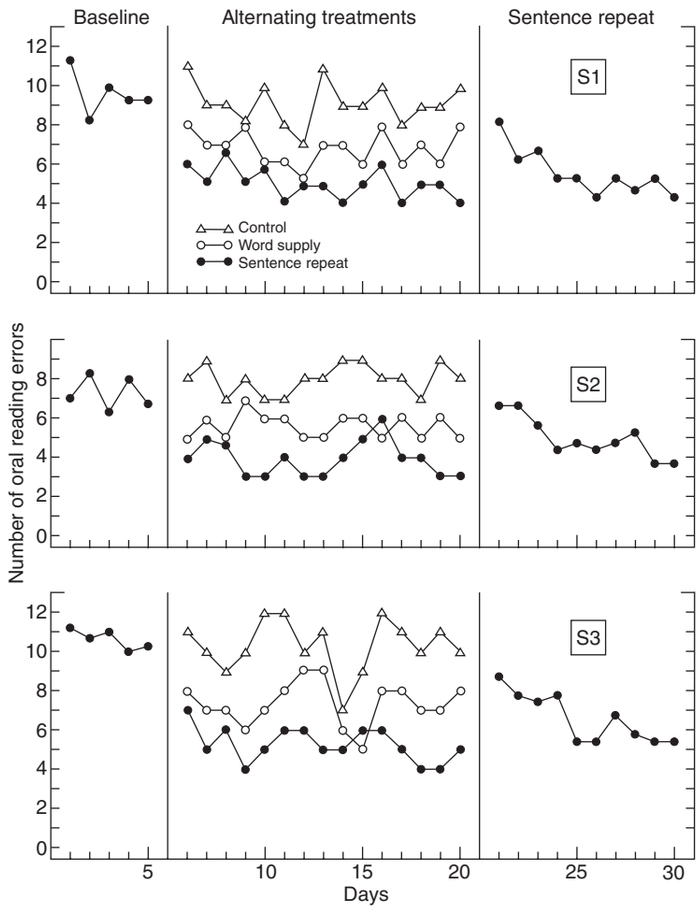


FIGURE 5–24
Research use of an alternating treatments design demonstrating a functional relationship.
Note: From "Effects of Two Error-Correction Procedures on Oral Reading Errors," by N. Singh, 1990, *Behavior Modification*. Copyright 1990 by Sage Publications, Inc. Reprinted by permission.

Marcia Learns Sight Vocabulary

Mr. Hagan was a resource teacher for elementary students. He wanted one of his students, Marcia, to learn basic sight vocabulary on a first-grade level. He chose 15 words and established that Marcia's baseline rate for reading them was zero. Mr. Hagan then divided the words into three sets of five. One set he printed on cards accompanied by an audio recording tape that Marcia could use to hear the words pronounced. He assigned a peer tutor to work with Marcia on the second set, and the teacher worked with Marcia on the third set. Mr. Hagan recorded and graphed the number of words Marcia pronounced correctly each day for each set. Within a week, Marcia was pronouncing correctly the group of words learned with the peer tutor at a higher rate than either of the other sets. Mr. Hagan concluded that, for Marcia, peer tutoring was the most efficient way to learn sight vocabulary.

Advantages and Disadvantages

The alternating treatments design is an efficient way for teachers to answer one of the most important instructional questions: Which method is most likely to be successful with this student? Once clear fractionation appears, the teacher can select the most successful method using as few as three to five data points. One disadvantage is the necessity to institute a replication phase in order to establish a clear functional relation. However, this is likely to be of little practical importance to teachers.

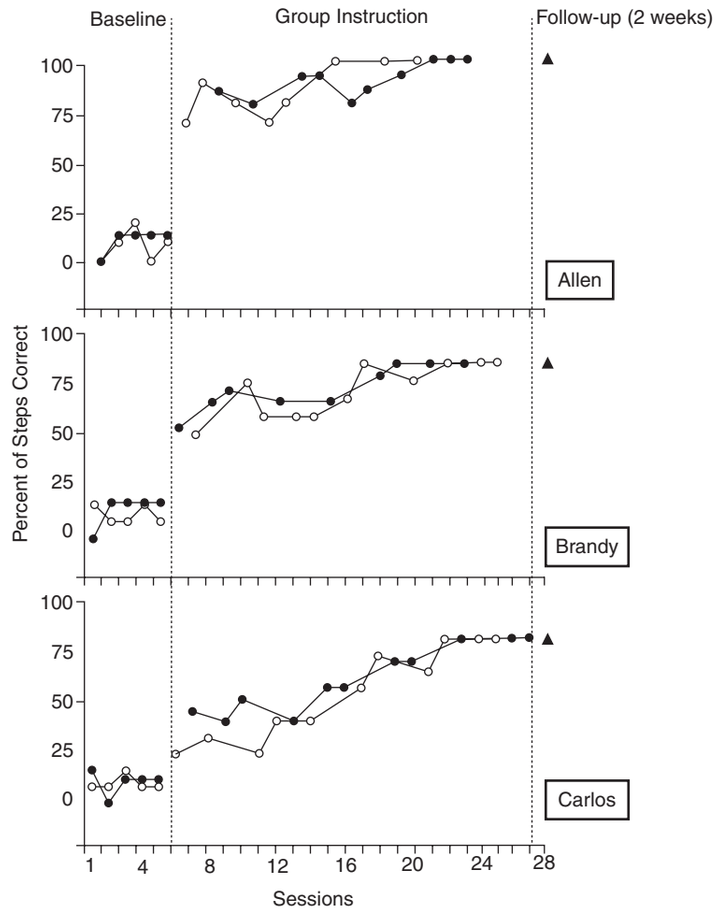


FIGURE 5–25
Use of an alternating treatments design.
Note: From "A Comparison of Static Picture Prompting and Video Prompting Simulation Strategies Using Group Instructional Procedures," by D. Cihak, P. Alberto, T. Taber–Doughty, R. Gama, 2006, *Focus on Autism and Other Developmental Disabilities, 21*, 89–99. Copyright 2006, by the Hammill Institute on Disability. Reprinted with permission.

CHANGING CONDITIONS DESIGN

The changing conditions design reflects reality—teachers keep trying different techniques until they find one that works.

A **changing conditions design** is used to investigate the effects of two or more treatments (independent variables) on the behavior of a student (dependent variable). Unlike the alternating treatments design, the treatments in this design are introduced sequentially. The changing conditions design is also referred to as a multiple treatments design, or an ABC design, because each new treatment phase is given an identifying letter (Cooper, 1981; Kazdin, 1982; Richards et al., 1999).

The design is useful for the teacher who finds it necessary to try a number of interventions before finding one that is successful with a particular student. The teacher is changing the conditions (for example, environmental conditions, instructional conditions, reinforcement conditions) under which the student is expected to perform the behavior.

IMPLEMENTATION

The first step in implementing a changing conditions design is to collect baseline data to assess the student's present level of performance. Once a stable baseline is established, the teacher can introduce the selected intervention and measure its effectiveness through

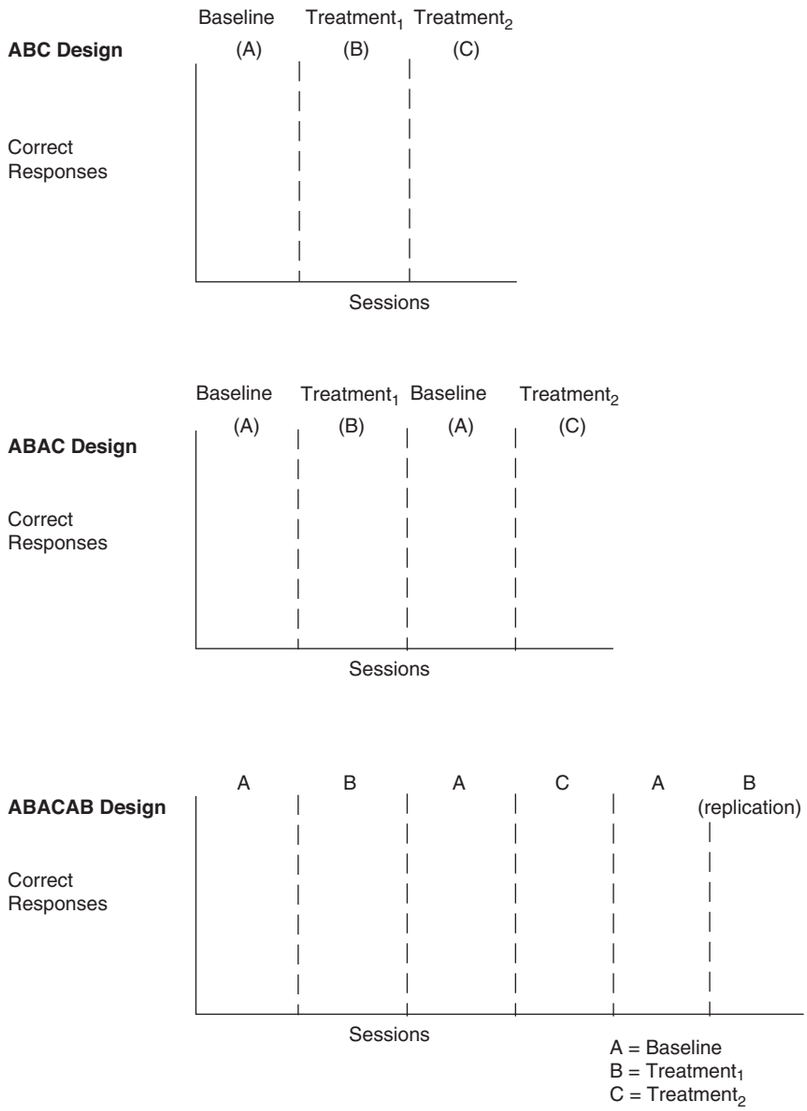


FIGURE 5–26 ***Variations of the changing conditions design.***

data collection. If the data for the first intervention do not demonstrate a change in the student's performance or if the change is not of a sufficient magnitude or in the desired direction, the teacher may design a second intervention. This second intervention can be either a complete change in strategy or a modification of the earlier intervention. This process of redesigning intervention conditions is repeated until the desired effect on the student's behavior is achieved. The teacher should expect to see some evidence of a change in the student's behavior within five intervention sessions.

The changing conditions design has three basic variations: (1) ABC, (2) ABAC, and (3) ABACAB (see Figure 5–26).

1. *ABC design:* The ABC design is used when the teacher is trying to judge the effectiveness between treatments, is trying to put together an instructional package that

will facilitate a student's performance, or is trying systematically to remove forms of assistance to bring a student to a more independent performance.

(a) Building an instructional package: Starting from a student's current performance, the teacher implements an intervention. If the student's performance does not respond or does not respond sufficiently, new strategies are successively or cumulatively added until the student's performance meets criterion. This format is compatible with current models of response-to-intervention (RTI). As each piece is added to the instructional package, a new phase is identified. This design is simply an extended AB design. As in an AB design, there is no replication of the effect of interventions and there can be no assumption of a functional relation. Smith (1979) was trying to improve the oral reading of a student with learning disabilities. Following baseline, three cumulative phases were employed. First the teacher used teacher modeling. When the change in the student's performance was not sufficient, a correction procedure was added to the modeling. When this combined strategy still did not produce sufficient change, previewing was added. This package of three strategies was successful.

(b) Fading assistance: The teacher systematically reduces the amount of assistance being provided to a student in order to identify the least amount needed for ongoing successful performance. Each reductive change is considered a new phase. Reductive changes might include reducing the intensity of antecedents, such as providing a student who is learning to print letters to trace in the first phase, densely spaced dots to connect in the second phase, more sparsely spaced dots in subsequent phases, and eventually to writing in the presence of just a line on the paper. Other changes might include reducing the amount of reinforcement or the frequency (schedule) of the delivery of reinforcement. Reducing the number of components of an instructional package is also an example of fading assistance. If, in order to be successful in writing a paragraph, a student initially needs to be given the topic, a picture depicting the topic, guidance through a verbal description of the picture, and a topic sentence, the teacher systematically removes each until the student is able to write a paragraph given just a topic. Examples of the use of fading in this design may be seen in studies by Ardoin, Martens, and Wolfe (1999), Boyle and Hughes (1994), and Oliver, Hall, Hales, Murphy, and Watts (1998).

2. *ABAC design:* In this design the teacher's implementation of two or more interventions is separated by additional baseline conditions: baseline, treatment 1, baseline, treatment 2, and so forth. The treatments may be completely different or variations of one other. Separating the treatments by intervening baseline conditions prevents one treatment's continuing to affect the student's behavior while another treatment is being used, thus providing a clear picture of the effect of each of the treatments. This design may be seen as a variation of an ABAB design. It is not, however, considered definitive in establishing a functional relation (Richards et al., 1999; Tawney & Gast, 1984). The study by Handen, Parrish, McClung, Kerwin, and Evans (1992) presented a variation of this design. As illustrated in Figure 5–27, two different interventions were compared; guided compliance was assessed versus time-out. Note the use of an intervening baseline separating the two treatments. Crozier and Tincani (2005) initially tried the use of social stories to reduce talk-outs. When talk-outs were not reduced to zero occurrences they tried the combined use of social stories plus prompts (i.e., "What's the rule for talking in school?"). This variation has an element

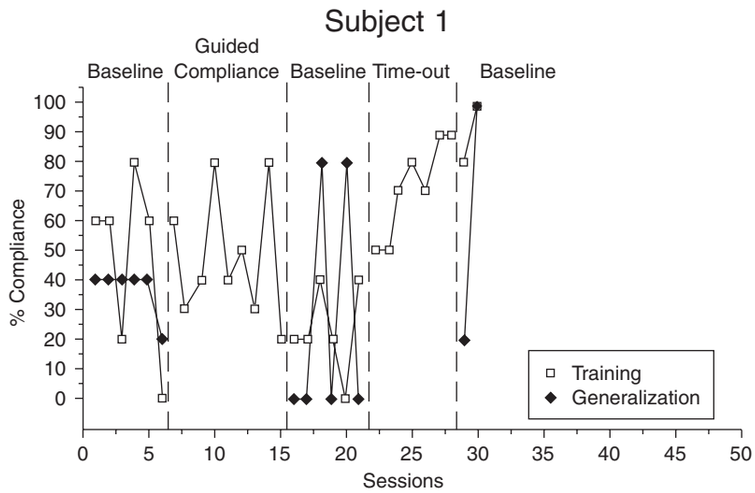


FIGURE 5–27
Use of a changing conditions design.
Note: From "Using Guided Compliance Versus Time Out to Promote Child Compliance: A Preliminary Comparative Analysis in an Analogue Context," by B. Handen, J. Parrish, T. McClung, M. Kerwin, & L. Evans, 1992, *Research in Developmental Disabilities.* Copyright 1992 with kind permission from Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington OX5 1GB, UK.

of building an intervention package. In Figure 5–27, note again the use of an intervening baseline separating the two treatments.

3. *ABACAB design:* The data resulting from an ABC or ABAC design do not allow for a determination of a functional relation between the dependent variable and any of the independent variables. As is the case with an AB design, the data can give only an indication of the effectiveness of a particular intervention. The design can, however, be refined in order to demonstrate a functional relation. To assess the presence of a functional relation, there should be a replication of the effect of the intervention; therefore, following phases for each of the potential treatments, the one whose data indicate it is most successful is reimplemented after another baseline condition. If the treatment is successful again, this is a replication of its effect, and therefore a functional relation is demonstrated. This design may also be seen as a variation of an ABAB design. A variation of this design may be seen in the studies by Cole, Montgomery, Wilson, and Milan (2000), who used an ABACAD, and Falk, Dunlap, and Kern (1996), who used an ABABCB.

Graphic Display

The format for the changing conditions design is similar to that of the previous designs. A baseline phase is followed by the intervention phases, with a dashed vertical line separating the sessions and data associated with each specific intervention. Figure 5–26 illustrates the three basic formats: ABC, ABAC, and ABACAB.

Research Application

Smith (1979) used a changing conditions design to measure the effect of a number of teaching conditions on a 12-year-old boy's oral reading (see Figure 5–28). The dependent variable measured was the number of words read orally by the student (per minute) and the number of reading errors made. (Note that the numbers in the circles represent the average number of words read per phase. The numbers in the rectangles represent

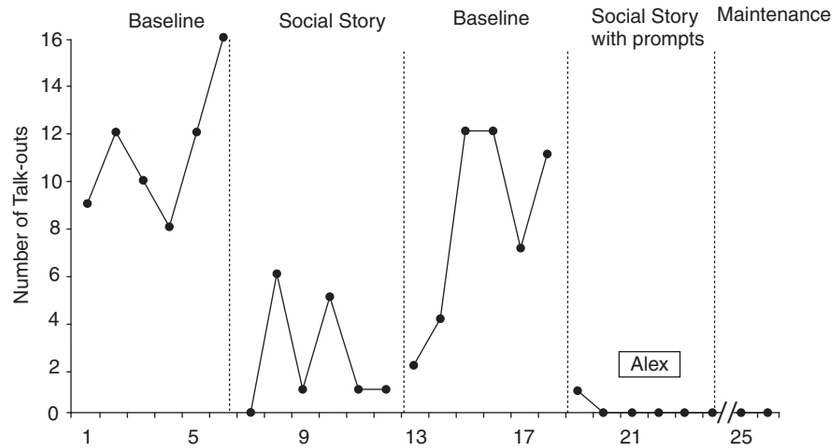


FIGURE 5–28
Use of a changing conditions design.
Note: From "Using a Modified Social Story to Decrease Disruptive Behavior of a Child with Autism," by S. Crozier & M. Tincani, 2005, *Focus on Autism and Other Developmental Disabilities, 20*, 150–157. Copyright 2005, by PRO-ED, Inc. Reprinted with permission.

the average number of errors per phase.) Figure 5–28 presents the data recorded under each of the following conditions:

1. *Baseline.* John was asked to read from his book.
2. *Modeling.* The teacher read the first page of a new story from the child's text. John was then asked to read orally.
3. *Modeling plus correction.* The previous condition was altered by adding a correction procedure. The teacher corrected John when he made an error and offered the correct word if he did not know it.
4. *Modeling plus previewing and correction.* After the teacher read, John reread the same passage and continued reading until the instructional time (5 minutes) elapsed. The correction procedure remained in effect.
5. *Follow-up.* Baseline conditions were reinstituted.

Handen et al. (1992) used a changing conditions design with repeated baseline to examine the relative efficacy of guided compliance and time-out as a method of increasing adherence to adult requests by young children with mild mental retardation. The following conditions (phases) were included in the study. In all phases, each child was presented with 10 requests (five targeted and five generalization probes). Compliance within 10 seconds was followed by praise. Figure 5–29 presents the data recorded for one of the children.

1. *Baseline.* Noncompliance was ignored and compliance praised.
2. *Guided compliance.* Noncompliance within 10 seconds resulted in the adult's guiding the child to complete the task using hand-over-hand assistance. Praise was withheld if assistance was provided.
3. *Baseline condition.*
4. *Time-out.* Noncompliance within 10 seconds resulted in placement of the child in a chair facing a corner of the room for 30 seconds (the child was held gently in the chair if he or she refused to remain seated).
5. *Baseline condition.*

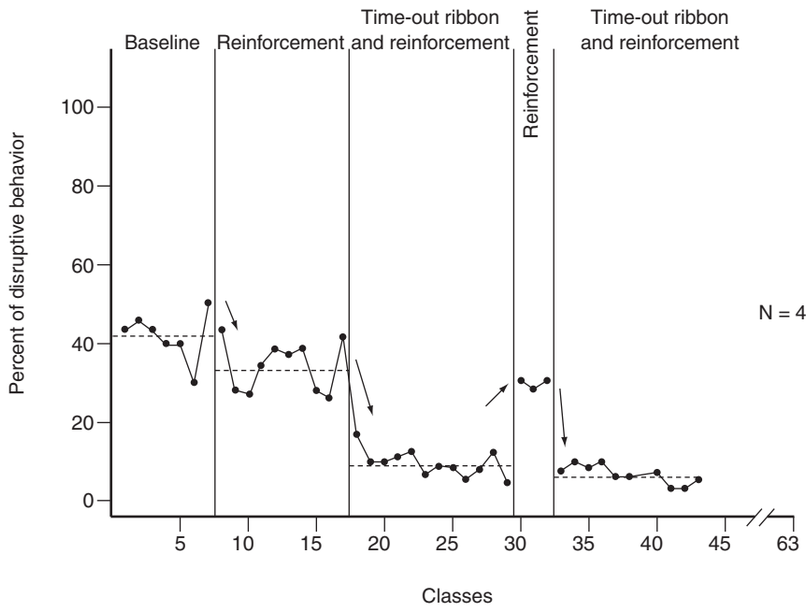


FIGURE 5–29
Graph for visual inspection of data.
Note: From "The Timeout Ribbon: A Nonexclusionary Timeout Procedure," by R. Foxx & S. Shapiro, 1978, *Journal of Applied Behavior Analysis.* Copyright 1978 by The Society for the Experimental Analysis of Behavior. Reprinted by permission.

Teaching Application

Here's how a changing conditions design can be used in teaching.

Roberta Learns to Shoot Baskets

Mr. Woods was recently hired to teach physical education at an elementary school. When he arrived at work, Mr. Woods was approached by the special education teacher, Ms. Jones. She was concerned about Roberta, a student with physical disabilities who would be in Mr. Woods's gym class. Roberta, who used a wheelchair, had difficulty with eye–hand coordination. Ms. Jones hoped the student could learn to throw a basketball. Learning to play basketball would provide coordination training and a valuable leisure skill for Roberta.

Baseline.

Mr. Woods agreed that the basketball skill seemed appropriate.

Mr. Woods decided to use a systematic approach to instruction. He asked Roberta to throw the basketball 20 times to see how often she could place the ball through a lowered hoop. This procedure was followed for five gym periods with no additional instruction until a baseline performance rate was determined. Mr. Woods then decided to use a modeling technique. He showed Roberta how to throw the ball and asked her to imitate him. Very little improvement was noted in five class periods. Mr. Woods met with the special education teacher to determine what could be done.

First condition.

Ms. Jones carefully reviewed all the data and suggested changing the conditions. She explained that a change in intervention seemed necessary and that a modeling procedure could be used in combination with keeping score on a chart.

Second condition.

Mr. Woods agreed to try this. In 2 weeks, Roberta showed improvement but still missed more baskets than she hit. A final condition was implemented using modeling, scorekeeping, and a correction procedure. Mr. Woods now showed Roberta how to throw, recorded her score, and showed her exactly what she did wrong when she missed. This combination of procedures resulted in Roberta's being able to throw a basketball through a hoop 15 out of 20 times. A suggestion was made to Roberta's parents that a hoop be constructed at her home so that she could enjoy her new skill after school.

Third condition.

Advantages and Disadvantages

Changing conditions is a teaching design.

The changing conditions design with a single baseline allows the teacher to compare the effects of a number of interventions on student behavior. Although no functional relation can be established, recording data in this format allows the teacher to monitor the effects of various procedures on student behavior. The teacher should be aware, however, that what she may be seeing is the cumulative effects of the interventions rather than the effects of any one intervention in isolation. Individual analysis of the effects of the interventions can be made using the repeated baselines format of the changing conditions design. The teacher who records data systematically in a changing conditions design will have a record of the student's progress and a good indication of what procedures are effective with that student. The six single-subject designs we described were the AB, ABAB (reversal), changing criterion, multiple baseline, alternating treatments, and changing conditions. A summary of the uses, formats, and types of questions answered by each of these designs is presented in Table 5–2.

Evaluating Single-Subject Designs

Analysis of Results

The purpose of using applied behavior analysis procedures in the classroom is to achieve, and verify, meaningful changes in a student's behavior. The effectiveness of an intervention can be judged against both an experimental criterion and a clinical criterion. The experimental criterion verifies that an independent variable (an intervention) was responsible for the change in the dependent variable (a behavior). Single-subject designs demonstrating within-subject replications of effect satisfy this criterion (Baer, Wolf, & Risley, 1968; Barlow & Hersen, 1984; Kazdin, 1998; Poling et al., 1995; Richards et al., 1999).

The *clinical criterion* is a judgment as to whether the results of the teacher's intervention are "large enough to be of practical value or have impact on the everyday lives of those who receive the intervention, as well as those in contact with them" (Kazdin, 2001, p. 153). For example, the teacher should ask herself whether it is truly meaningful to increase a student's grade from a D− to a D (Baer et al., 1968), or to decrease a student's self-injurious behavior from 100 to 50 instances per hour (Kazdin, 2001), or to reduce a student's off-task behavior in the special education class while it remains high in a general educational class. The teacher should ask if the student's behavior has decreased sufficiently so as to no longer interfere with other students' learning, or with the ability of her family to carry out its activities at home and in the community.

A third criterion for evaluating the outcome of an intervention is its *social validity.* Those involved with the student's educational program should be concerned about and evaluate the social acceptability of an intervention program and of its outcome (Kazdin, 1977, 2001; Wolf, 1978). Social validity will be discussed at length in Chapter 12.

Visual Analysis of Graphs

Intervention effects in applied behavior analysis are usually evaluated through *visual analysis* of the graph displaying the plotted data points of the various phases (conditions). Certain characteristics of the data paths within and across phases are examined in order to judge the effectiveness of the intervention. These characteristics include the *mean* of the data points in the phase, the *levels* of performance from one phase to the next, the *trend* in performance across phases, the *percentage* of data that overlap in

TABLE 5–2
Summary of single-subject research designs.

Design	Use	Format	Example Questions
AB	To document changes in behavior during baseline and intervention. Does not allow for determination of a functional relationship—lacks a replication of the effect of the independent variable (intervention) on the dependent variable (behavior).	*Two Phases* 1. Baseline 2. Intervention	1. Will Sam's mastery of sight words increase when I use a time delay procedure for instruction? 2. Will Sam's call-out behavior decrease when I reinforce hand raising with tokens?
ABAB Reversal	To determine if a functional relationship exists between an independent variable and a dependent variable by replicating the baseline and intervention phases.	*Four Phases* 1. Baseline 2. Intervention 3. Return to baseline 4. Return intervention	1. Could the number of words Sam writes in a paragraph increase due to use of a point system? 2. Could Sam's off-task behaviors decrease due to use of a self-recording procedure?
Changing Criterion	To increase or decrease a behavior in systematic increments towards a terminal criterion. Allows for determination of a functional relationship if performance level matches the continually changing interim criteria.	Baseline plus an intervention phase for each interim criterion towards the objective; e.g., 10 phases of interim criteria raising 5 words per phase until criterion of 50 words.	1. Can I use a time delay procedure to systematically increase the pool of Sam's sight words to a criterion of 100 words? 2. Can I use token reinforcers to systematically decrease the number of times Sam runs in the hall during change of class to a criterion of no occurrences?
Multiple Baseline	To determine if a functional relationship exists between an independent variable and a dependent variable by assessing replication/generalization across: a) behaviors, b) individuals, or c) settings.	Baseline and staggered intervention phase for each replication/generalization, e.g., across behaviors: baseline and intervention for Sara's callouts *and* then out-of-seat behaviors; across students: baseline and intervention for cursing by Bob *and* then Ted; across settings: baseline and intervention in the resource class *and* then consumer math class.	1. a) Will the use of reinforcement of Linda's hand raising result in a decrease in the number of occurrences of both call-outs and out-of-seat without permission? b) Will the use of a learning strategy such as content mastery increase Linda's completion of American history assignments and biology assignments? 2. a) Will the use of points to earn the opportunity to be a team captain result in decrease in cursing by Bob, Ted, and Linda? b) Will the use of a calculator increase the accuracy of grocery purchasing by both Bob and Ted? 3. a) Will the use of self-recording result in a decrease in the occurrences of out-of-seat by Linda in the resource room, consumer math class, and music class? b) Will the use of tokens increase the number of math problems completed by Linda in both the resource class and consumer math class?

Continued

TABLE 5–2 (continued)

Design	Use	Format	Example Questions
Alternating Treatments	To determine which of two or more independent variables is more effective for increasing or decreasing occurrences of a dependent variable. Has the ability to determine the existence of a functional relation by replicating use of the more effective independent variable in an additional phase.	*Three Phases* 1. Baseline 2. Intervention phase in which each independent variable is applied on alternating days, or alternating sessions during the same day. 3. A functional relation may be determined by replication of the more effective intervention with the content taught by the less effective; or at the time period in which the less effective intervention was used.	1. Will the use of number line or counting chips increase Jane's accuracy in addition? 2. Will the use of earning points or losing points prove more effective in decreasing Jane's off-task behavior?
Changing Conditions Multiple Treatments ABC	To determine which of two or more independent variables is more effective for increasing or decreasing occurrences of a dependent variable. Functional relationship may be determined by replication of more effective independent variable following an additional baseline.	Multiple phases: e.g., Baseline First independent variable Baseline Second independent variable Baseline Possible replication of more effective independent variable.	1. Will oral and written practice increase Jane's spelling accuracy on test, or will oral practice alone be just as effective, or will written practice alone be just as effective? 2. Will point loss or point loss plus verbal reprimand be more effective in decreasing Jane's tardiness to classes?

adjacent phases, and the *rapidity* of behavior change within phases (Cooper et al., 2007; Kazdin, 1998, 2001; Kennedy, 2005; Richards et al., 1999).

1. Evaluation of changes in means focuses on the change in the average rate of student performance across the phases of a design. Within each phase, the mean (average) of the data points is determined and may be indicated on the graph by drawing a horizontal line corresponding to the value on the ordinate scale. Visual inspection of the relation of these means will help determine if the intervention resulted in consistent and meaningful changes in the behavior in the desired direction of change. In Figure 5–29, Foxx and Shapiro (1978) supplied such indicators of means. The viewer can easily see the relative position of the students' disruptive behavior across the various design phases.

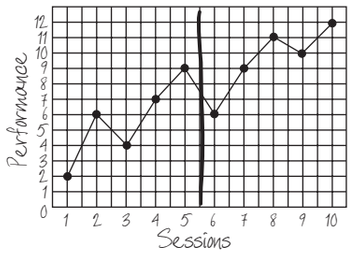
2. Evaluation of the level of performance refers to the magnitude and direction of the change in student performance from the end of one phase to the beginning of the next phase. "When a large change in level occurs immediately after the introduction of a new condition, the level change is considered abrupt, which is indicative of a powerful or effective intervention" (Tawney & Gast, 1984, p. 162). Tawney and Gast suggested the following steps to determine and evaluate a level change between two adjacent conditions: (1) identify the ordinate value of the last data point of the first condition and the first data point value of the second condition, (2) subtract the smallest value from the largest, and (3) note whether the change in level is in an improving or decaying direction (p. 162). In Figure 5–29, the arrows have been added to indicate level changes.

3. Evaluation of a trend in performance focuses on systematic and consistent increases or decreases in performance. Data trends are most often evaluated using a procedure known as the *quarter-intersect method* (White & Liberty, 1976). Evaluation of trends is based on lines of progress developed from the median value of the data points in each phase. The use of a trend line increases the reliability of visual analysis among people looking at a graph (Bailey, 1984; Ottenbacher, 1993; Ottenbacher & Cusick, 1991). This is of particular importance as teams of teachers, students, parents, and other concerned individuals review student data to assess progress and make decisions about future instruction or intervention. Steps for computing lines of progress are illustrated in Figure 5–30. Trend lines can provide (1) an indication of the direction of behavior change in the past and (2) a prediction of the direction of behavior change in the future. This information can help the teacher determine whether to change the intervention.

   Taking this process one step further will yield a split-middle line of progress (White & Haring, 1980). This line of progress is drawn so that an equal number of data points fall on and above the line as fall on and below the line. As illustrated in Figure 5–31, if the data points do not naturally fall in such a pattern, the line is redrawn higher or lower, parallel to the original line, until the balance of data points is equal.

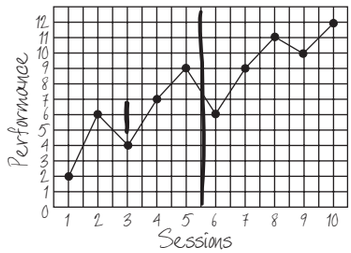
4. Evaluation of the percentage of overlap of data plotted for performance (ordinate values) across contiguous conditions provides an indication of the impact of an intervention on behavior. This is referred to as *effect size* and is used as a measure of effectiveness of the intervention (Kromrey & Foster-Johnson, 1996). Percent of overlap is calculated by "(1) determining the range of data point values of the first condition, (2) counting the number of data points plotted in the second condition, (3) counting the number of data points of the second condition that fall within the range of values of the first condition, and (4) dividing the number of data points

FIGURE 5–30 ***Steps for computing lines of progress.***



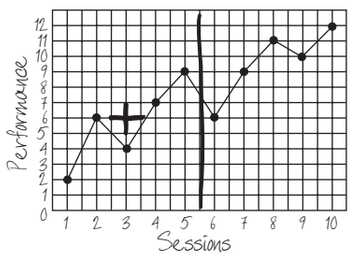
1. Divide the number of data points in half by drawing a vertical line down the graph.

In this example, there are 10 data points: therefore, the line is drawn between sessions 5 and 6. If there had been an odd number of data points, this would have been drawn through a session point.



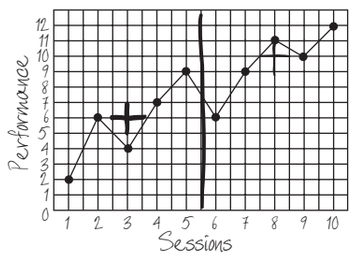
2. On the left half of the graph, find the midsession and draw a vertical line.

In this example, there are five data points; therefore, the line is drawn at session 3. If there had been an even number of sessions, this line would have been drawn between two session points.



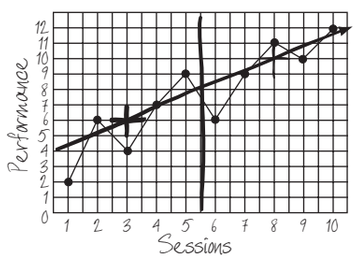
3. On the left half of the graph, find the mid-performance point and draw a horizontal line.

In this example, the data point at performance value 6 is the midperformance point because there are two data points below it and two data data points above it. If there had been an even number of data points, this line would have been drawn between the two media points.



4. Repeat steps 2 and 3 on the right half of the graph.

In this example, session 8 is the midsession, and the data point at performance value 10 is the midperformance point.



5. Draw a line connecting the intersections of both halves of the graph. This is the trend line for the data.

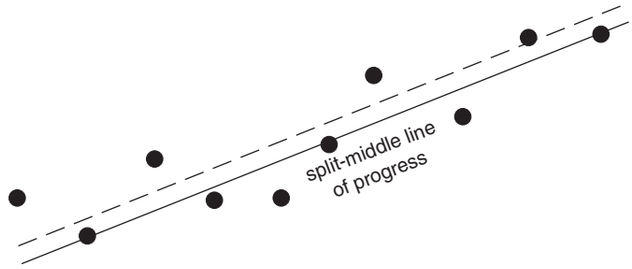


FIGURE 5–31
Split-middle line of progress.
Note: From *Exceptional Teaching* (p. 118), by O. White & N. Haring, 1980, Columbus, OH: Merrill. Reproduced by permission.

that fall within the range of the first condition by the total number of data points of the second condition and multiplying this number by 100. In general, the lower the percentage of overlap, the greater the impact the intervention has on the target behavior" (Tawney & Gast, 1984, p. 164).

For example, in Figure 5–29, the range of data values during baseline (phase 1) is 32 to 50. In the reinforcement-only condition (phase 2), 6 of 10 data points fall within the same data value range as the baseline, yielding a 60% overlap. However, the percentage of overlap between phase 2 and phase 3 is 0%. These percentages of change indicate that the use of a time-out ribbon and reinforcement had a much greater impact on the disruptive behavior than did reinforcement alone.

5. Evaluation of the rapidity of the behavior change (sometimes called the latency of behavior change) refers to the length of time between the onset or termination of one phase and changes in performance. The sooner the change occurs after the experimental conditions have been altered (that is, after implementation or withdrawal of the intervention), the clearer the intervention effect (Kazdin, 1998). Note that "rapidity of change is a difficult notation to specify because it is a joint function of changes in level and slope (trend). A marked change in level and in slope usually reflects a rapid change" (Kazdin, 1982, p. 316).

Visual analysis is often quick and effective and is relatively easy to learn (Poling et al., 1995). This makes it useful to the teacher trying to make instructional and behavior management decisions in the classroom. The use of visual analysis encourages ongoing evaluation as data are collected and phases change, rather than reliance on pre- and postintervention data. This facilitates data-based decision making for educational programming.

Problematic in the use of visual analysis is a lack of concrete decision rules for determining whether a particular demonstration shows or fails to show a reliable effect (Kazdin, 1998). The components of visual analysis do not have agreed-upon operationalized criteria based in the research literature. Each teacher or researcher sets the standard for a component as he or she uses it. Therefore, visual analysis may be seen as subjective and open to inconsistent application by an individual across sets of student data, or across different individuals reviewing student data. Confidence in conclusions based on visual inspection may be increased by increasing reliable use of the various components. The reliability may be increased by: (1) teacher training and repeated opportunities for use; (2) interpreting student performance data with a consistently applied standard; and (3) two or more trained individuals independently reviewing the data and drawing conclusions that can be compared (Richards et al., 1999). In special education there is at least annual interpretation and review of student data by the teacher and the IEP team. This presents an opportunity for review of data interpretation and collaboration in setting standards for data-based decision making.

It is noted that evaluation resulting from visual analysis reveals only intervention results that have a strong and reliable effect on behavior—it may miss consistent but subtle or weak behavior change caused by some interventions. However, it may be considered a benefit for classroom use that visual analysis is more likely to identify independent variables that produce strong or socially significant results. The usual purpose of intervention is to obtain immediate and strong treatment effects. If obtained, such effects are "quite evident from visual inspection" (Kazdin, 2001, p. 150). When single-subject designs are used in classroom decision making, clinical and social validity are important criteria. The clinical criterion is a judgment as to whether the results of the teacher's intervention are large enough to be of practical value and impact on the learning or behavior of the student. They are likely to be socially valid based on the functional change in the student's performance and the social acceptance of the student.

Although visual inspection is useful, convenient, and basically reliable for identifying or verifying strong intervention effects for decision making in the classroom, educational and behavioral researchers may choose to explore statistical evaluation of single-subject data as a companion to, or comparison with, the results of visual analysis (Richards et al., 1999). This may be the case when there is concern for generalization across populations, or when seeking intervention effects so subtle as not to be clinically significant but which further research might be able to make more significant or more consistent. Kazdin (1976) offered three reasons for use of statistical techniques: (1) distinguishing subtle effects from chance occurrence, (2) analyzing the effect of a treatment procedure when a stable baseline cannot be established, and (3) assessing treatment effects in environments that lack control. Information about advanced uses of visual inspection and statistical evaluation with single-subject designs can be found in Barlow and Hersen (1984), Cooper et al. (2007), Franklin, Allison, and Gorman (1996), Kazdin (1982, 1998), Kennedy (2005), Richards et al. (1999), and Tawney and Gast (1984).

Action Research and Single-Subject Design Tools

Action research is any systematic inquiry conducted by teachers and other educational professionals in teaching/learning environments to gather information and reflect upon how their school operates, how they teach, or how well their students learn. Information is gathered with goals including effecting positive changes in the classroom and school environments, and improving student outcomes (Mills, 2003, p. 5). Action research encourages teachers to be participant researchers to gather information to share with the educational team. This information allows immediate analysis of instructional and behavior management issues and is used to develop the next step(s) in programming. Teachers draw from research design tools that will describe what they are seeing in order to analyze and develop solutions and thereby improve their practice.

Action research is considered a naturalistic approach to research. Methods are considered naturalistic when they occur in natural settings (e.g., classrooms) with relatively little interruption of the normal flow of events. For the most part they are nonexperimental. Naturalistic researchers are not interested in manipulating or controlling the situation, and they are not interested in studying interventions to discover functional relations. When the purpose of research is to understand what is rather than to study that which has been manipulated and controlled, naturalistic research methods are appropriate (Arhar, Holly, & Kasten, 2001, p. 36). Fundamentally, action research is nonexperimental and descriptive, whereas single-subject research is experimental and seeks to identify functional relations resulting from the manipulation of variables.

Components of Action Research

There is general agreement on the basic steps of an action research study: (1) identify an area of focus or concern, (2) collect data for documentation, (3) analyze and interpret data, and (4) share the information with others and develop an action plan (Arhar et al., 2001; Schoen & Nolen, 2004; Stringer, 2004; Mills, 2003).

There is an eclectic array of data collection procedures available. As noted by Mills (2003), action research uses elements of quantitative (e.g., comparison of standard scores) and qualitative research methods. However, the literature emphasizes the data collection tools of qualitative research. These include use of observation, interviews, questionnaires, checklists, rating scales, focus groups, records, physical products, scatter plots, fieldnotes, anecdotal records, videotape, audiotapes, and photographs (Arhar et al., 2001; Mills, 2003; Stringer, 2004). Action researchers often use frequency counts or percentages to describe the extent of behaviors. Arhar et al. suggest the importance of frequencies in capturing the scope of a behavior as reflected in questions such as How often does "this" occur? How often does "this" occur in comparison to "that"? Does it occur constantly and evenly? Does it occur periodically or in waves?" (p. 201). In order to assess consistency and patterns of behavior, line graphs and bar graphs are used for organizing and visual display of these data.

Single-subject Design Parallels and Contributions

In order to broaden the number and range of tools available to teachers when planning action research, contributions from single-subject research should be considered. They can provide data collection techniques and descriptive graphing tools for quick and easy implementation in the classroom. These can be used by the teacher as a participant and can provide objective data. Some tools of single-subject research parallel existing recommended procedures, and some must be added.

a. *Parallel procedures:* Three areas of single-subject data collection make use of the methodologies similar to those in the action research literature. First, permanent product recording makes use of written records, videotapes, audiotapes, photographs, and physical results of behavior. These data are transformed into frequency or percentage data. Second, we make use of anecdotal recording in order to describe and analyze chains of behavior. As presented in Figure 3–2, we make use of a strategy for structuring these observations to assist in analysis. Third, the methods and data sheets for collecting event data discussed in Chapter 3 will help a teacher structure the collection of frequency and percentage data. Additionally, Chapter 6 includes a discussion of questionnaires, scales, and an alternative scatter plot procedure ABA practitioners use in association with functional assessment.

b. *Procedures to add:* Several single-subject designs are appropriate for the descriptive purposes of action research. The AB design displays and monitors behavior once a plan of action is implemented. The ABC design monitors the effects on behavior of adding components to an instructional package. The alternating treatments design, using only the first two phases, allows display and monitoring of the effects of two interventions. Each of these designs allows one to see if the behavior being examined changed with the implementation of an intervention. However, with these designs there is no manipulation of an intervention and therefore no assessment of a functional relation.

Example of an Action Research Study

The study conducted by Schoen and Nolen (2004) is an example of the use of some single-subject research tools as part of action research. The teacher and the team addressed the behavior of a sixth grader with learning disabilities. He participated in general education and special education classes. His acting-out behaviors were causing him to be off-task and resulting in poor academic engagement and lack of academic success. This therefore was the identified focus of concern. Several kinds of data were collected to inform decision making and action planning: (a) Focused observation in the form of antecedent, behavior, consequence analysis (ABC analysis) was conducted over a 5-day period. This analysis identified specific behavior patterns that included slamming materials, yelling at teacher/peers, muttering under his breath, storming out of the room, destroying his work, and tuning out (head down on the desk). (b) Interviews were conducted with the student, the special education teacher, and the social worker. (c) A literature review of various theories and strategies was conducted. These data were shared with the educational team and an action planned was developed. A package of peer modeling, a self-management checklist, and positive reinforcement was put in place during reading, math, and transitions. The self-management checklist had the student assess his behavior with the following questions: Did I yell out? Did I stay on-task? Did I act respectfully to other students and teachers? Did I use proper outlets to calm down? The team chose for analysis the total minutes off task displayed on a graph showing baseline data and data during use of the action plan—thus the AB design, as shown in Figure 5–5. This type of graph allowed the team members and the student to monitor the decrease in the number of minutes off task due to his acting-out behavior.

Summary

This summary serves as a rationale and an answer to the question: "Of what use is this to me?"

By best practice and legal mandate, data-based demonstrations of learning are required as evidence of effective instruction and of a quality education. Applied behavior analysis provides tools to meet these accountability requirements. Chapter 3 introduced methods of collecting data that provide the raw material for discussions of effectiveness. This chapter introduced single-subject designs as ways to organize the gathering and display of data. The design routinely used in the classroom is the AB design because it is a direct reflection of common classroom practice. It does not require restructuring teaching sessions. The graph of the AB design provides an uncomplicated visual format that can be used by teacher, student, parent, and supervisor to monitor, interpret, and assess learning. The other single-subject designs have specific capabilities and therefore may be used less frequently. In various ways, each design provides a database for quick, student-specific decision making. Table 5–2 summarizes the use of each and the questions they attempt to answer.

After the AB design, the changing criterion design is the most direct reflection of how a teacher manages instruction. Teachers regularly break down a goal that requires a large amount of learning into manageable units. Teaching manageable units of content, one at a time, in sequence, is the graphic picture resulting from organizing instruction and data collection within a changing criterion design. Another decision often required is which of two or more strategies will result in the most effective and efficient learning. The alternating use of strategies within the format of the alternating treatments design provides a data-based answer, usually by the end of 1 week. Given a little more time, an answer to the question can also be provided by using the ABC

design. Variations of this design are used more commonly to evaluate the combination of several strategies in an instructional package. The multiple baseline design has gained popularity as inclusive policies are put in place in schools. Of particular interest is the multiple baseline across settings that allows tracking the effectiveness of an intervention across general education, special education, community, and home settings. The reversal design allows quick and unobtrusive evaluation of an intervention on a classroom problem that you do not want to allow to progress from being a nuisance to a spreading classroom management problem. The reversal design is appropriate for issues such as out of seat, off task, and not doing homework. Later in this text, this design is used when developing behavior management plans resulting from functional behavior assessment.

Aspects of instruction and behavior management are being continually researched and evaluated. This is especially true in ABA, which has a culture of data-based decision making. Research brings to the classroom extensions of current strategies and evaluation of proposed strategies. Teachers must be able to answer questions like: Is what I am doing still the best practice? Do the suggestions being made by colleagues, supervisors, and parents have a basis in data-based research? If educators are to be lifelong learners, they must be able to access the information provided in professional research journals. In order to read these journals, one must be literate in the type of research being published. From an ABA perspective, one must be able to read research conducted with single-subject designs. Most often found in the research journals are reversal and multiple baseline designs. The reversal design (ABAB) is frequently used because it is most powerfully able to demonstrate a functional relation between a behavior and an intervention due to the controlled application and removal of a strategy (Kazdin, 1982). Multiple baseline designs are found frequently because they build immediate replication and therefore depth of experience with the intervention that allows broader and more confident statements of applicability to other students, behaviors, or settings.

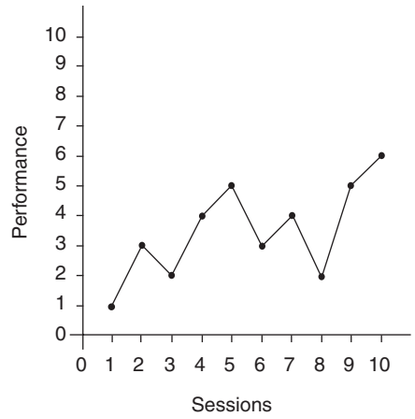
Key Terms

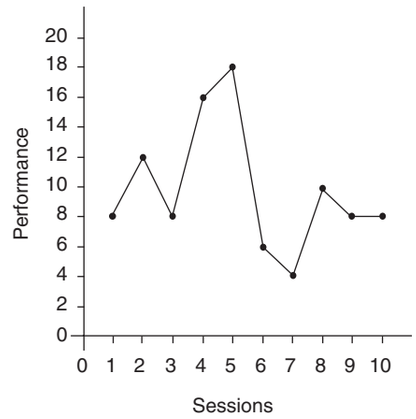
variable
dependent variable
independent variable
functional relation
group designs
single-subject designs
enhanced functioning
baseline data
trend
AB design
reversal design
ABAB design
changing criterion design
multiple baseline design
alternating treatments design
changing conditions design

Discussion Questions

1. Baseline data for Craig's self-injurious behavior indicate a mean occurrence of 17 instances per 40-minute observation period. What change in his behavior would be clinically significant (as demonstrated by enhanced functioning)?
2. Which single-subject design might a teacher use to systematically introduce and teach 30 community sight words?
3. During 3 weeks of multiplication instruction, probes of Alison's performance indicated that she still could not multiply. Her teacher wants to determine which of two alternative approaches to teaching multiplication would be most effective for Alison. Select two instructional methods. Select an appropriate single-subject design and outline the steps the teacher should follow to make this determination.

4. Outline a procedure associated with a single-subject design that would demonstrate the generalization of an intervention across settings in a high school.
5. Draw lines of progress on the two sets of data graphed below.





6. Many of the studies that appear in professional journals use "embedded" designs. That is, one single-subject design is embedded within another. This is illustrated in the following graph. (a) Identify the components of the multiple baseline within the graph; (b) identify the components of the reversal within the graph; and (c) identify the elements demonstrating a functional relation.