CHAPTER

Defining and Assessing Learning

Concept: People who assess learning must make inferences from observing performance during practice and tests

APPLICATION

Any professional involved in motor skills instruction typically has to provide some type of assessment to determine whether or not the student or patient has learned what the professional has taught. The following two situations, common in physical education and rehabilitation settings, provide examples of the importance of assessing learning.

Suppose you are a physical educator teaching a tennis unit. If you are teaching your students to serve, how do you determine if they are actually learning what you are teaching them? What will you look for to assess their progress in learning to serve? How can you be certain that what you are observing is the result of learning and not just luck?

Or suppose you are a physical therapist helping a stroke patient to learn to walk without support. What evidence will tell you that this patient is learning to do what you have taught him or her to do? What characteristics of the patient's performance will make you confident that the patient is learning this skill and will be able to walk without assistance at home as well as in the clinic?

The answers to these questions are important for effective professional practice in any setting in which people need to learn motor skills. As you think about possible answers to these questions, consider two important characteristics of learning that you need to take into account whenever you assess skill learning. First, we do not directly observe learning; we directly observe behavior. Second, because of this, we must make inferences about learning from the behavior we observe. Any learning assessment procedure must incorporate these two critical characteristics of learning. In the discussion that follows, we will address these points by first, establishing a definition of learning, and then discussing several learning assessment procedures.

DISCUSSION

In any discussion about the assessment of learning, we need to keep two important terms distinct: *performance* and *learning*. This distinction helps us establish an appropriate definition for the term *learning;* it also helps us consider appropriate conditions under which we should observe performance so that we can make valid inferences about learning.

Performance Distinguished from Learning

Simply put, **performance** is *observable behavior.* If you observe a person walking down a corridor, you

A CLOSER LOOK

The Terms "Performance" and "Learning"

Performance	Learning
• Observable behavior	• Inferred from performance
• Temporary	• Relatively permanent
• May not be due to practice	• Due to practice
• May be influenced by performance variables	• *Not* influenced by performance variables

are observing him or her perform the skill of walking. Similarly, if you observe a person hitting a baseball, you are observing a performance of the skill of hitting a ball. When used in this way, the term *performance* refers to the execution of a skill at a specific time and in a specific situation. *Learning,* on the other hand, cannot be observed directly, but can only be inferred from characteristics of a person's performance.

Before considering a more formal definition for learning, think about how often we make inferences about people's internal states based on what we observe them doing. For example, when someone smiles (an observable behavior), we infer that he or she is happy. When someone cries, we infer that he or she is sad, or perhaps very happy. When a person yawns, we assume that the person is tired. In each of these situations, certain characteristics about the individual's behavior are the basis for our making a particular inference about some internal state we cannot observe directly. However, because we must base our inference on observed behavior, it is possible for us to make an incorrect inference. For example, if a student sitting beside you in class yawns during the lecture, you might infer from that behavior that the person is tired because of lack of sleep the night before. However, it may be that he or she is bored.

Learning defined. We will use the following general definition for the term **learning:** *a change in the capability of a person to perform a skill that must be inferred from a relatively permanent improvement in performance as a result of practice or experience.* It is important to note from this definition that the person has increased his or her *capability, or potential,* to perform that skill. Whether or not the person actually performs the skill in a way that is consistent with this potential will depend on the presence of what are known as *performance variables.* These include factors that can affect a person's performance but not the degree of learning the person has achieved. Some examples include the alertness of the person, the anxiety created by the situation, the uniqueness of the setting, fatigue, and so on. As a result, it is critical that methods used to assess learning take factors such as these into account to allow accurate inferences about learning.

General Performance Characteristics of Skill Learning

We generally observe four performance characteristics as skill learning takes place.

Improvement. First, *performance of the skill shows improvement over a period of time.* This means that

performance the behavioral act of performing a skill at a specific time and in a specific situation.

learning a change in the capability of a person to perform a skill; it must be inferred from a relatively permanent improvement in performance as a result of practice or experience.

the person performs at a higher level of skill at some later time than at some previous time. It is important to note here that learning is not necessarily limited to improvement in performance. There are cases in which practice results in bad habits, which in turn result in the observed performance's failure to show improvement. In fact, performance actually may become worse as practice continues. But because this text is concerned with skill acquisition, we will focus on learning as it involves improvement in performance.

Consistency. Second, as learning progresses, *performance becomes increasingly more consistent.* This means that from one performance attempt to another, a person's performance characteristics should become more similar. Early in learning, performance is typically quite variable from one attempt to another. Eventually, however, it becomes more consistent.

A related term here is *stability,* which was introduced in chapter 4. As performance consistency of a skill increases, certain behavioral characteristics of performance become more stable. This means that the acquired new behavior is not easily disrupted by minor changes in personal or environmental characteristics.

Persistence. The third general performance characteristic we observe during learning is this: *the improved performance capability is marked by an increasing amount of persistence.* This means that as the person progresses in learning the skill, the improved performance capability lasts over increasing periods of time. A person who has learned a skill should be able to demonstrate the improved level of performance today, tomorrow, next week, and so on. However, because of some forgetting or other factors, the person may not achieve the same performance level on each of these occasions as he or she did at the end of the practice time devoted to the skill. The persistence characteristic relates to the emphasis in our definition of learning on a *relatively permanent improvement* in performance.

Adaptability. Finally, an important general characteristic of performance associated with skill learning is that *the improved performance is adaptable to a variety of performance context characteristics.* We never really perform a skill for which everything in the performance context is exactly the same each time. Something is different every time we perform a skill. The difference may be our own emotional state, the characteristics of the skill itself, an environmental difference such as a change in weather conditions, the place where we perform the skill, and so on. Thus, successful skill performance requires adaptability to changes in personal, task, and/or environmental characteristics. The degree of adaptability required depends on the skill and the performance situation. As a person progresses in learning a skill, his or her capability to perform the skill successfully in these changed circumstances also increases. Later in this book, we will explore some instruction and practice condition characteristics that can influence how well a person adapts to these various situations.

Assessing Learning by Observing Practice Performance

One way we can assess learning is to record levels of a performance measure during the period of time a person practices a skill. A common way to do this is to illustrate performance graphically in the form of a **performance curve.** This is a plot of the level achieved on the performance measure for each time period, which may be time in seconds or minutes, a trial, a series of trials, a day, etc. For any performance curve, the levels of the performance measure are always on the Y-axis (vertical axis), and the time over which the performance is measured is on the X-axis (horizontal axis).

Performance curves for outcome measures. We can graphically describe performance by developing a performance curve for an outcome measure of performance. An example is shown in figure 11.1, which depicts one person's practice of a complex pursuit tracking task. The task required the person to track, or follow the movement of, a cursor on a

A CLOSER LOOK

Examples of Motor Skill Performance Adaptability Demands

Closed Skills	
• *Hitting a sand wedge in golf*	• from wet sand, dry sand, etc. • from various locations in the sand trap • to various pin locations on the green • when shot has various implications for score
• *Shooting free throws in basketball*	• one- and two-shot free throws at various times of the game • one-and-one shot situations at various times of the game • with various crowd conditions (e.g., quiet, loud, visible behind the basket) • various types of backboards
• *Walking*	• on various types of surfaces • in various settings (e.g., home, mall, sidewalk) • while carrying various types of objects • alone or while carrying on a conversation with a friend
Open Skills	
• *Hitting a baseball/softball*	• various types, speeds, and locations of pitches • various ball-and-strike counts • various people-on-base situations with various numbers of outs • left-handed and right-handed pitchers
• *Catching a ball*	• balls that are different shapes, weights, sizes, etc. • various speeds and directions • in the air, on the ground • with one or two hands
• *Driving a car*	• various sizes of cars • various street and highway conditions • with or without passengers • various weather conditions

computer monitor by moving the mouse on a tabletop. The goal was to track the cursor as closely as possible in both time and space. Each trial lasted about 15 sec. The outcome measure of performance was the root-mean-squared error (RMSE).

Notice that in this graph we can readily observe two of the four behavioral characteristics associated with learning. First, *improvement* is evident by the general direction of the curve. From the first to the last trial, the curve follows a general downward trend (indicating decreasing error). Second, we can also see *increased performance consistency* in this graph. The indicator of this performance characteristic is performance on adjacent trials. According to figure 11.1, this person showed a high degree of inconsistency early in practice but became slightly

performance curve a line graph describing performance in which the level of achievement of a performance measure is plotted for a specific sequence of time (e.g., sec, min, days) or trials; the units of the performance measure are on the Y-axis (vertical axis) and the time units or trials are on the X-axis (horizontal axis).

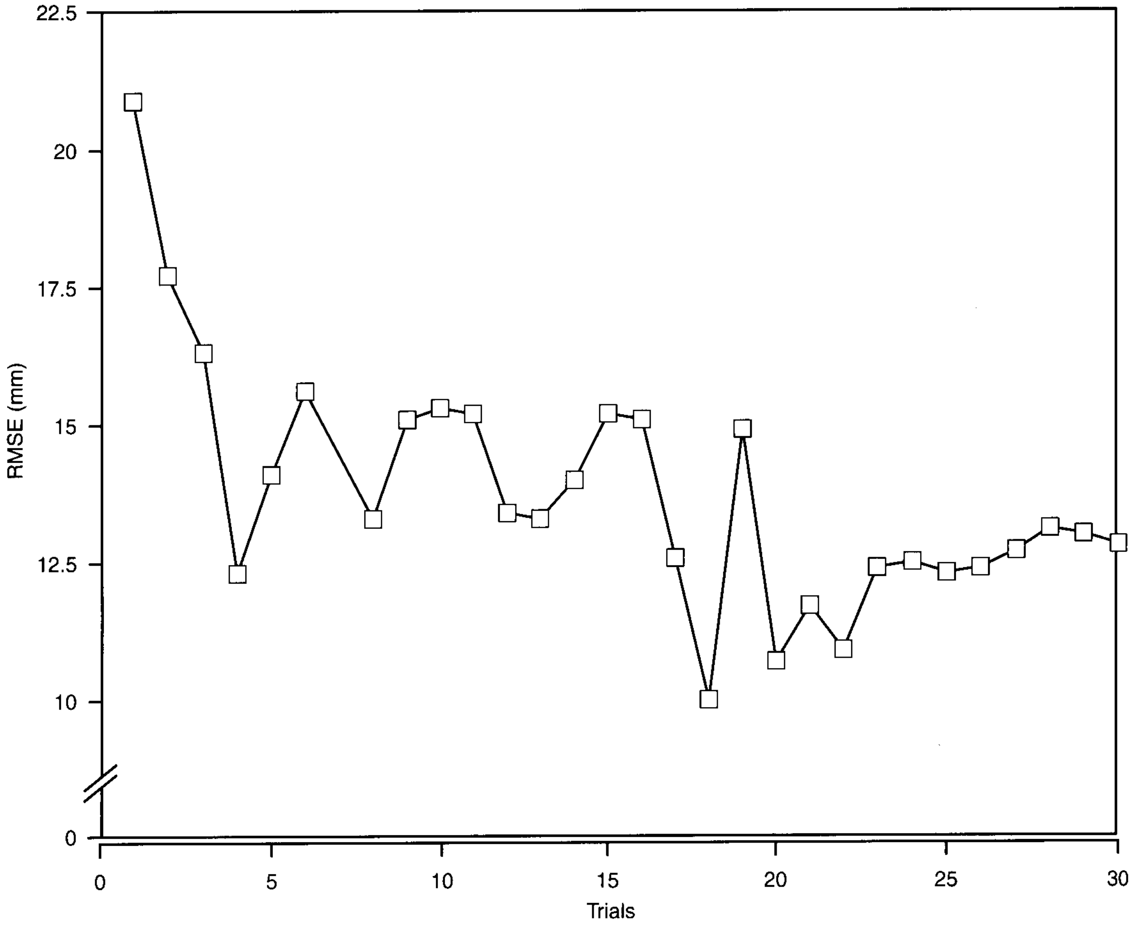

FIGURE 11.1 Performance curve for one person learning a pursuit tracking task. The performance measure is the root-mean-squared error (RMSE) for each trial.

more consistent from one trial to the next toward the end of practice. The expectation would be that the person would increase this consistency with additional practice trials.

When a person is acquiring a new skill, the performance curve for an outcome measure typically will follow one of *four general trends* from the beginning to the end of the practice period for a skill. This period of time may be represented as a certain number of trials, hours, days, etc. The trends are represented by the four different shapes of curves in figure 11.2. Curve (a) is a *linear curve,* or a straight line. This indicates proportional performance increases over time; that is, each unit of increase on the horizontal axis (e.g., one trial) results in a proportional increase on the vertical axis (e.g., one second). Curve (b) is a *negatively accelerated curve,* which indicates that a large amount of improvement occurred early in practice, with smaller amounts of improvement later. This curve represents the classic power function curve of skill learning, which we will discuss in some detail in chapter 12 as a characteristic of the power law of practice. Curve (c) is the inverse of curve

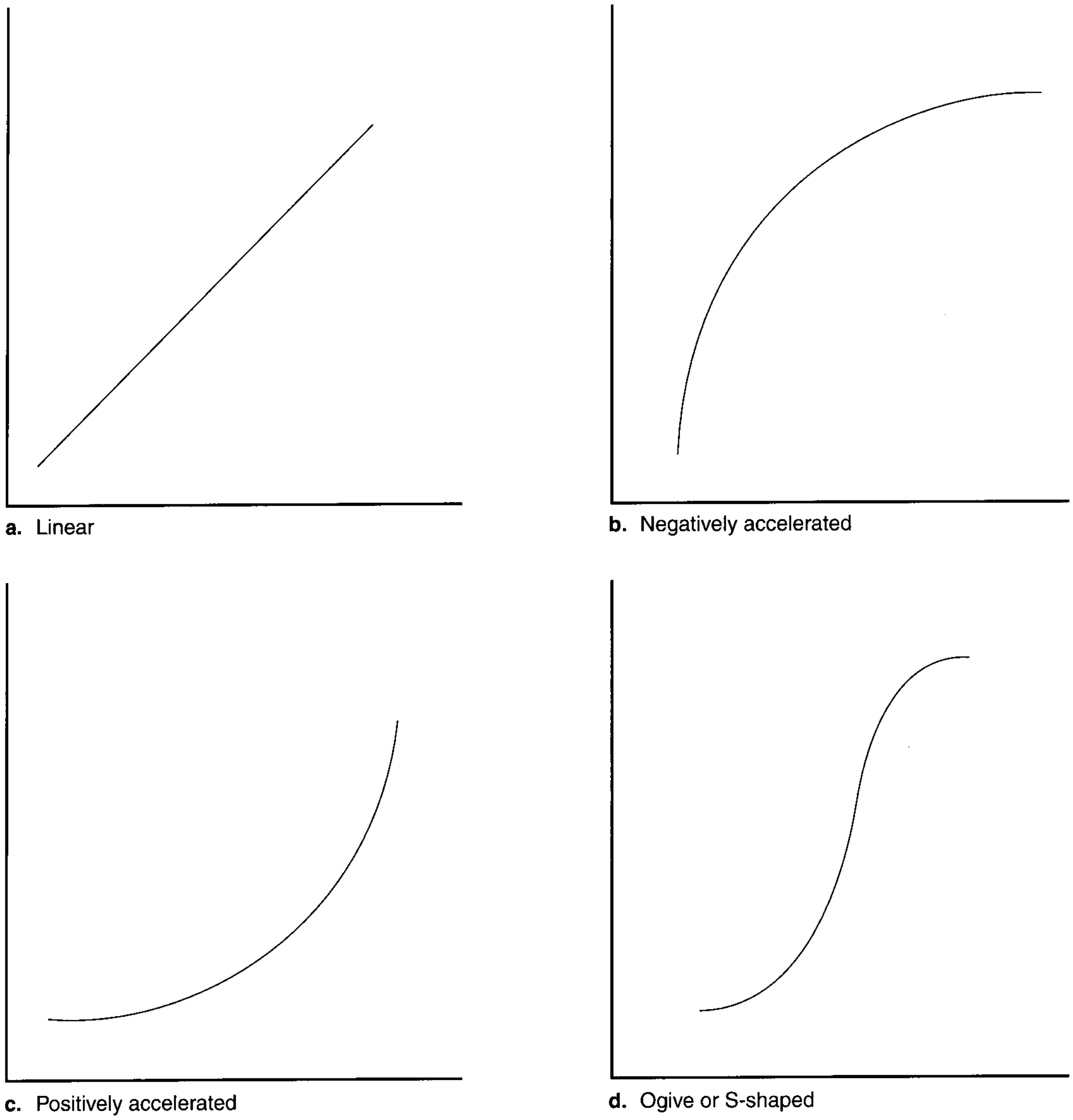

FIGURE 11.2 Four general types of performance curves.

(b) and is called a *positively accelerated curve.* This curve indicates slight performance gain early in practice, but a substantial increase later in practice. Curve (d) is a combination of all three curves, and is called an *ogive* or *S-shaped curve.*

It is important to note that each curve in figure 11.2 shows better performance as the curve slopes upward. There are, however, instances in which the slope of the curve in a downward direction indicates performance improvement. This occurs when the performance measure is one for which a decrease in the performance level means better performance. Measures involving error or time (such as absolute error and reaction time) follow this characteristic as performance is improving when the amount of error or time decreases. In such cases, the directions of the performance curves would be opposite to those just described, although the shapes of the curves would be the same.

The four curves presented in figure 11.2 are hypothetically smoothed to illustrate general patterns of performance curves. Typically, performance curves developed for individuals are not smooth but erratic, like the one in figure 11.1. However, there are various statistical procedures that can be used for curve smoothing when the reporting of research results warrants it. Finally, various individual, instructional, and motor skill characteristics can influence the type of curve that will characterize a person's performance as he or she learns a skill. You will learn about several of these characteristics in various chapters of this textbook.

Performance curves for kinematic measures. When we use performance production measures, such as kinematics, we cannot always develop performance curves like the one in figure 11.1. This is because a kinematic measure typically does not lend itself to being represented by one number value for each trial. Kinematic measures involve performance over time *within* a trial. It is important to include this time component in the graphic representation of a kinematic measure.

To assess improvement and consistency in performance for a series of practice trials, researchers commonly show one performance curve graph for each trial. To show improvement and consistency changes, they depict a representative sample of trials from different stages of practice.

We can see an example of this approach to kinematic measures in figure 11.3. The task that generated these measures is commonly used in motor learning research. It required participants in an experiment to move a lever on a tabletop to produce the criterion movement pattern shown at the top of this figure. Each participant observed the criterion movement on a computer monitor. The four graphs located below the criterion movement pattern represent the performance of one person for 800 trials. To provide a more representative picture of performance, the researchers analyzed practice trials in blocks of ten trials each. To represent performance changes during practice, figure 11.3 shows four blocks of trials, each representing a different segment of the 800-trial session. Each graph shows two performance characteristics. One is the person's average pattern drawn for the block of trials; this is indicated by the solid line (mean). The second is the variability of the patterns drawn for that same block of trials; this is indicated by the dashed lines (SD, or standard deviation).

To determine *improvement in performance,* compare the early to the later practice trials by examining how the shape of the person's produced pattern corresponds to the shape of the criterion pattern. As the person practiced more, the produced pattern became more like the criterion pattern. In fact, in trials 751 through 760, the participant was making a pattern almost identical to the criterion pattern.

To assess *changes in consistency,* compare how far the standard deviation lines are from the mean pattern for each block of trials. For trials 1 through 10, notice how far the standard deviation lines are from the mean. This shows a large amount of trial-to-trial variability. However, for trials 751 through 760, these lines are much closer to the mean, indicating that the person more consistently produced the same pattern on each trial of that block of trials.

Assessing Learning by Retention Tests

Another means of inferring learning from performance examines *the persistence characteristic of improved performance* due to practicing a skill. A common means of assessing this characteristic is to administer a retention test. You have been experiencing this approach to assessing learning since you began school. Teachers regularly give tests that cover units of instruction. They use these **retention tests** to determine how much you know, or have retained from your study. The teacher makes an inference concerning how much you

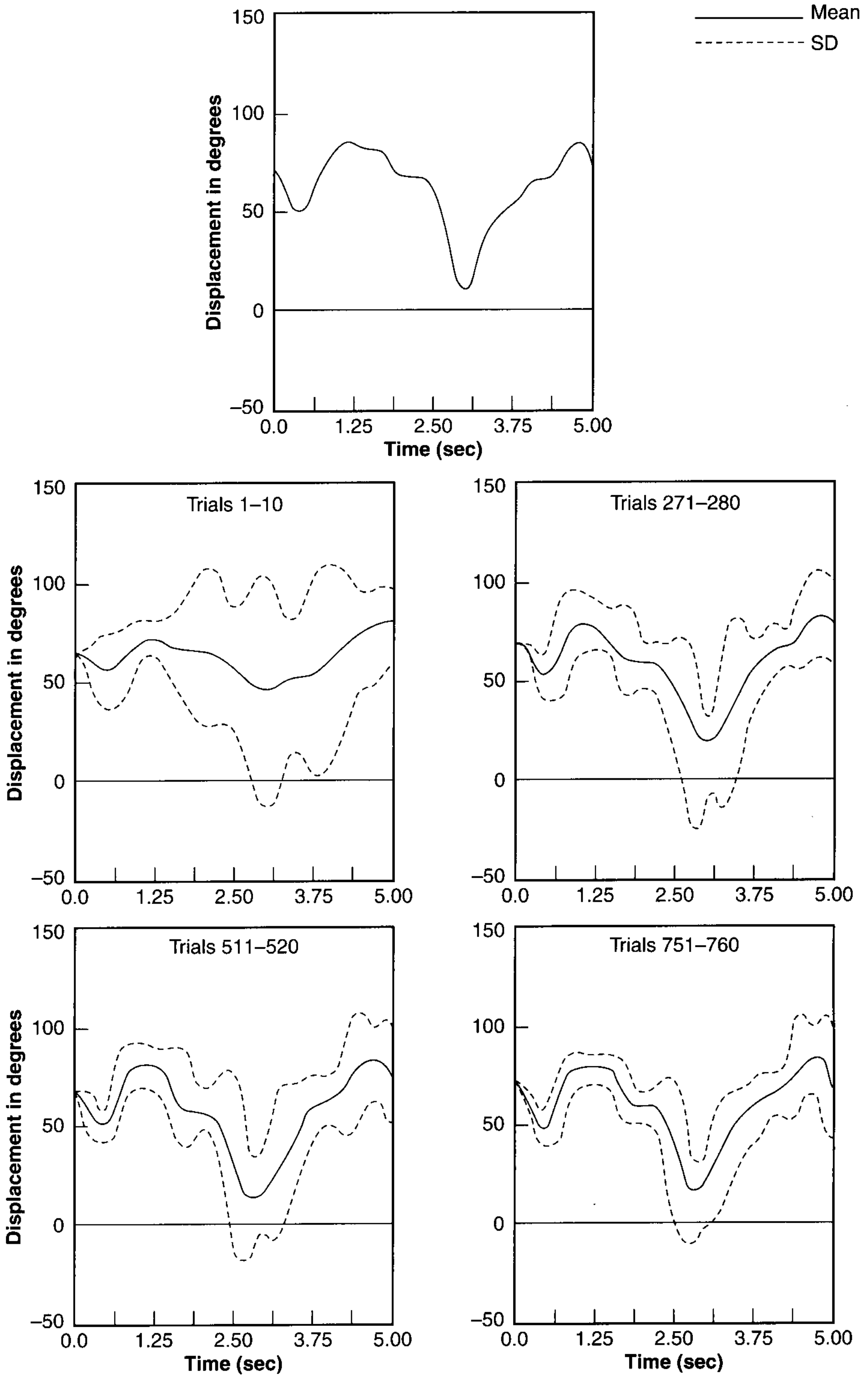

FIGURE 11.3 Results of an experiment by Marteniuk and Romanow showing changes in performance accuracy (displacement) on a tracking task at different practice trial blocks for one participant. The graph at the top shows the criterion pathway for the tracking task. [From Marteniuk, R. G., & Romanow, S. K. E. (1983). Human movement organization and learning as revealed by variability of movement, use of kinematic information, and fourier analysis. In R. A. Magill (Ed.), *Memory and control of action.*]

have learned about a particular unit of study on the basis of your test performance.

The typical way to administer a retention test in a motor skill situation is to have people perform the skill they have been practicing after a period of time during which they have not actually practiced the skill. The purpose is to determine the degree of permanence or persistence of the performance level achieved during practice; having a period of time with no practice allows this type of assessment. The actual length of time between the end of practice and the test is arbitrary. But the amount of time should be sufficiently long to allow the influence of any performance variables to dissipate to determine what was learned during practice. The critical assessment is the difference between the person's performance level on the first practice day and that on the test. If there is a significant improvement between these two periods of time, then you can be confident that learning has occurred.

Assessing Learning by Transfer Tests

The third means of inferring learning examines the *adaptability aspect of performance changes* related to learning. This assessment method involves using **transfer tests,** which are tests involving some novel situation, so that people must adapt the skill they have been practicing to the characteristics of this new situation. Two types of novel situations are especially interesting. One is a new context in which the people must perform the skill; the other is a novel variation of the skill itself.

Novel context characteristics. Test administrators can use various kinds of context changes in transfer tests. One characteristic they can change is the *availability of augmented feedback,* which is the performance information a person receives from some external source. For example, in many practice situations, the person receives augmented feedback in the form of verbal information about what he or she is doing correctly or incorrectly. If you were assessing learning to discover how well the person can rely on his or her own resources to perform the skill, then your requirement that the person perform without augmented feedback availability would be a useful context change for the transfer test. It is important to note that some researchers refer to a test that involves this type of context change as a retention rather than a transfer test, because the practiced skill is performed during the test.

Another context characteristic a test administrator can change is the *physical environment* in which a person performs. This is especially effective for a learning situation in which the goal is to enable a person to perform in locations other than those in which he or she has practiced. For example, if you are working in a clinic with a patient with a gait problem, you want that patient to be able to adapt to the environmental demands of his or her everyday world. Although performing well in the clinic is important, it is less important than performing well in the world in which the patient must function on a daily basis. Because of this need, the transfer test in which the physical environment resembles one in the everyday world is a valuable assessment instrument.

The third aspect of context that can be changed for a transfer test is the *personal characteristics* of the test taker as they relate to skill performance. Here, the focus is on how well a person can perform the skill while adapting to characteristics of himself or herself that were not present during practice. For example, suppose you know that the person will have to perform the skill in a stressful situation. A test requiring the person to perform the skill while emotionally stressed would provide a useful assessment of his or her capability to adapt to this situation.

Novel skill variations. Another aspect of adaptability related to skill learning is a person's capability to successfully perform a novel variation of a skill he or she has learned. This capability is common in our everyday experience. For example, no one has walked at all speeds at which it is possible to walk. Yet, we can speed up or slow down our walking gait with little difficulty. Similarly, we have not grasped and drunk from every type of cup or glass that exists in the world. Yet when we are confronted with some new cup, we adapt our movements quite well to the cup characteristics and successfully drink from it. These examples illustrate the importance to people of producing novel variations of skills. One of the ways to assess how well a person can do this is to use a transfer test that incorporates this movement adaptation characteristic.

Note that one of the ways we get people to produce a novel skill variation is to alter the performance context in some way so that they must adapt their movements to it. In this way, the transfer test designed to assess capability to produce novel skill variations resembles a transfer test designed to assess capability to adapt to novel performance context features. The difference is the learning assessment focus.

Assessing Learning from Coordination Dynamics

Another method of assessing learning involves the observation of the stabilities and transitions of the dynamics of movement coordination related to performing a skill. Proponents of this approach, which is gaining in popularity in learning research, assume that when a person begins to learn a new skill, he or she is not really learning something new, but is evolving a new spatial and temporal coordination pattern from an old one. When viewed from this perspective learning involves the transition between the initial pattern, represented by a preferred coordination mode the person uses when first attempting the new skill, to the establishment of the new coordination mode. *Stability and consistency* of the coordination pattern are important criteria for determining which coordination state (initial, transition, or new) characterizes the person's performance.

For example, a person who is learning handwriting experiences an initial state represented by the coordination characteristics of the upper arm, forearm and hand while engaged in handwriting at the beginning of practice. These characteristics make up the preferred spatial and temporal structure the person and the task itself impose on the limb, so the limb can produce movement approximating what is required. This initial stable state must be changed to a new stable state in which the person can produce fluent handwriting. Learning is the transition between these two states.

An example of this approach to assessing skill learning is an experiment by Lee, Swinnen, and Verschueren (1995). The task (see figure 11.4) required participants to learn a new bimanual coordination pattern. To perform the task, they simultaneously moved two levers toward and away from the body at the same rate (15 times in 15 sec). Their goal was to produce ellipses on the computer monitor. But to accomplish this, they had to coordinate the movement of their arms so that the right arm on each cycle was always 90 degrees out of phase with the left arm. The initial coordination pattern for the two arms for one participant is shown in figure 11.4 as the arm-to-arm displacement relationship demonstrated on the pretest on the first day of practice. The series of diagonal lines resulted when the person moved the arms in phase, in a motion resembling that of windshield wipers. The stability of this coordination pattern is indicated by the consistency of the fifteen diagonal lines produced during the pretest trial, and by the person's tendency to produce that same pattern on the pretest trial on day 2, after having performed sixty practice trials of the ellipse pattern on day 1.

By the end of day 3, this person had learned to produce the ellipse pattern. Evidence for this is the consistent production of fifteen ellipses in both the pretest and the posttest trials on day 3. However,

retention test a test of a practiced skill that a learner performs following an interval of time after practice has ceased.

transfer test a test in which a person performs a skill that is different from the skill he or she practiced, or performs the practiced skill in a context or situation different from the practice context or situation.

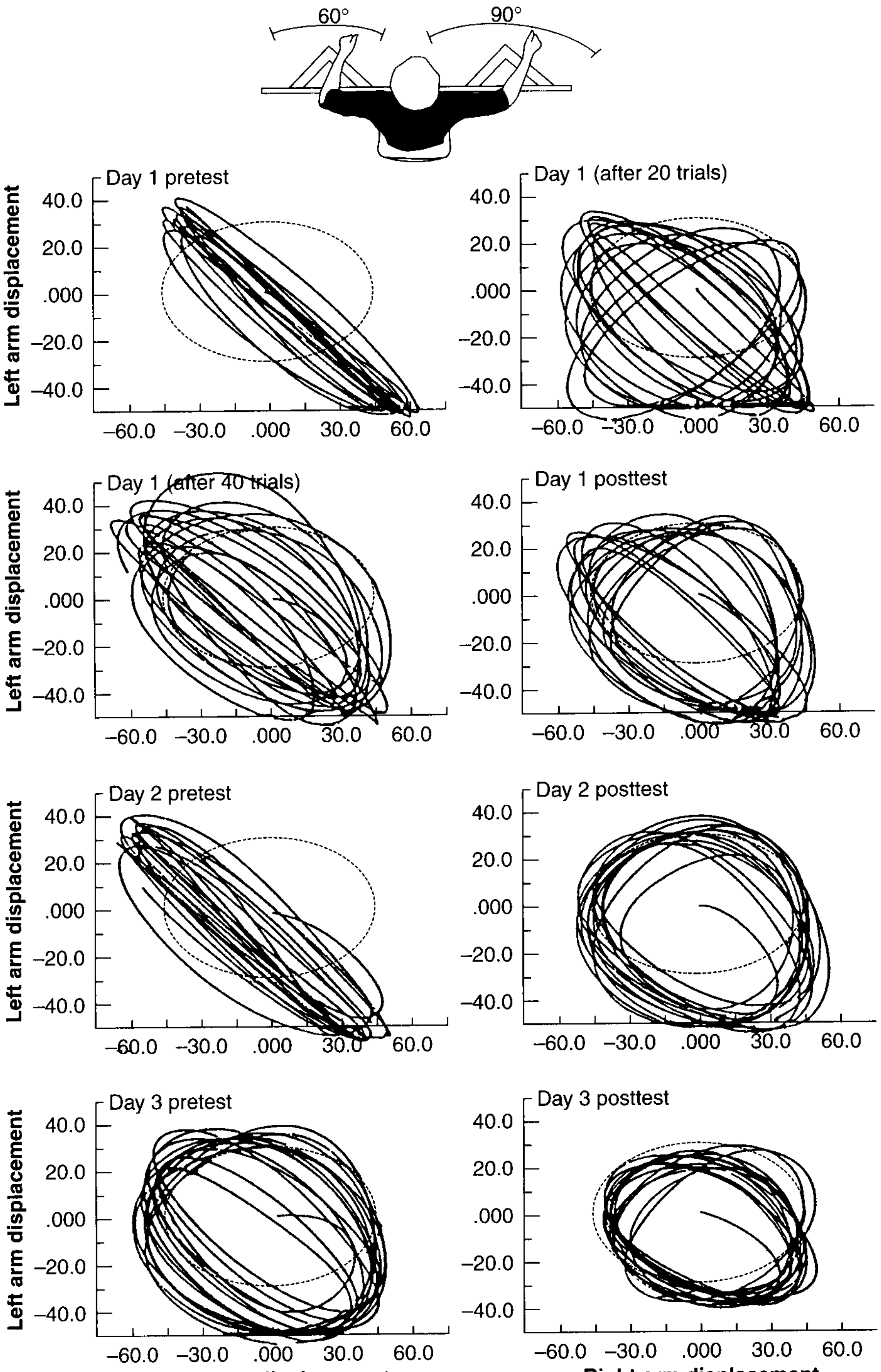

FIGURE 11.4 The task and results from the experiment by Lee, Swinnen, and Verschueren. The top panel shows the task, in which participants moved two levers to draw ellipses on the computer monitor (the dotted lines on each graph represent the goal ellipse pattern). The series of graphs shows the results as the left-arm × right-arm displacements of one person for the pretest and posttest (and some intermediate) trials for each of three practice days. [From Lee, T. D. et al. (1995). Relative phase alterations during bimanual skill acquisition. *Journal of Motor Behavior, 27*, 263–274. Reprinted with permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publications, 1319 Eighteenth Street NW, Washington, DC 20036–1802. Copyright © 1995].

notice the inst
many trials bet
terns (exhibited
posttest). This
tion between tw
process of learn

Practice Perfor
Misrepresent L

It may be misl
learning solely
practice. There
One is that the p
formance variab
this discussion a
inflate or depres
is that practice p
involves perform

Practice perform
estimate learnin
examples of vari
tice influences p
formance overest
One way to ove
retention or trans
son's practice per
then that person'
should demonstr
and not deviate t
mance at the end
performance sho
increased capabili

Performance plat
a skill, it is not u
ence a period of
seems to have st
some later time, im
This period of tim
be no further perf
as a **performance**

Examples of pe
to find in the mo
because most of thi

notice the instability of the performance in the many trials between the old and the new stable patterns (exhibited on the day 1 pretest and the day 3 posttest). This instability occurs during the transition between two stable states and characterizes the process of learning a new skill.

Practice Performance May Misrepresent Learning

It may be misleading to base an inference about learning solely on observed performance during practice. There are at least two reasons for this. One is that the practice situation may involve a performance variable, which was described earlier in this discussion as having the potential to artificially inflate or depress performance. The second reason is that practice performance may be misleading if it involves performance plateaus.

Practice performance may overestimate or underestimate learning. In this textbook, you will see examples of variables whose presence during practice influences performance in such a way that performance overestimates or underestimates learning. One way to overcome these problems is to use retention or transfer tests to assess learning. If a person's practice performance does represent learning, then that person's performance on a retention test should demonstrate the persistence characteristic and not deviate too much from his or her performance at the end of practice. Similarly, transfer test performance should demonstrate the person's increased capability to adapt to novel conditions.

Performance plateaus. Over the course of learning a skill, it is not uncommon for a person to experience a period of time during which improvement seems to have stopped. But for some reason, at some later time, improvement starts to occur again. This period of time during which there appears to be no further performance improvement is known as a **performance plateau.**

Examples of performance plateaus are difficult to find in the motor learning research literature because most of this research presents performance curves that represent the average for a group of participants. To find evidence of a performance plateau, individual participant's results are needed. An experiment reported by Franks and Wilberg (1982) is an example of this latter case, and it provides a good illustration of a performance plateau (figure 11.5). This graph shows one individual's performance on a complex tracking task for ten days, with 105 trials each day. Notice that this person showed consistent improvement for the first four days. Then, on days 5 through 7, performance improvement stopped. However, this was a temporary characteristic; performance began to improve again on day 8 and the improvement continued for the next two days. The steady-state performance on days 5 through 7 is a good example of a performance plateau.

The concept of a performance plateau has had a historical place in motor learning research. The first evidence of a plateau during skill learning is attributed to the work of Bryan and Harter (1897), who published their observations of new telegraphers learning Morse code. The authors noted steady improvement in the telegraphers' letters-per-minute speed for the first twenty weeks. But then a performance plateau occurred that lasted six weeks; this was followed by further performance improvement for the final twelve weeks. Since this early demonstration, researchers have been debating about whether a plateau is a real learning phenomenon or merely a temporary performance artifact (see Adams 1987 for the most recent review of plateau research). At present, most agree that plateaus are *performance rather than learning characteristics.* This means that plateaus may appear during the course of practice, but it appears that learning continues during these times.

performance plateau while learning a skill, a period of time in which the learner experiences no improvement after having experienced consistent improvement; typically, the learner then experiences further improvement with continued practice.

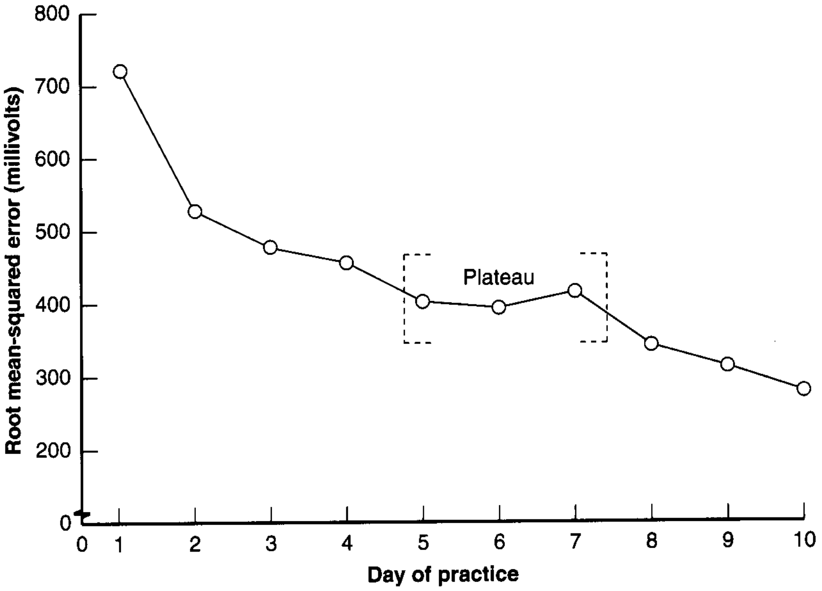

FIGURE 11.5 Results from the experiment by Franks and Wilberg showing the results from one participant performing the complex tracking task for ten days, with 105 trials per day. Notice the performance plateau for three days (days 5, 6, and 7) where performance leveled off before the subject showed improvement again. [From Franks, I. M., & Wilberg, R. B. (1982). The generation of movement patterns during the acquisition of a pursuit tracking task. *Human Movement Science, 1,* 251–272. Copyright © 1982 Elsevier/North-Holland, Amsterdam, The Netherlands. Reprinted by permission.]

There are several *reasons performance plateaus occur.* One is that the plateau represents a period of transition between two phases of acquiring certain aspects of a skill. During this transition, the person is developing a new strategy that the task requires to increase the level of performance already achieved. Consequently, no performance improvement occurs until the new strategy is successfully implemented. Other possible explanations for a performance plateau may be that it represents a period of poor motivation, a time of fatigue, or a lack of attention directed to an important aspect of a skill. Finally, it is possible the plateau may be due not to these performance characteristics but to limitations imposed by the performance measure. This is the case when the performance measure involves what are known as *ceiling* or *floor effects.* These effects occur when the performance measure will not permit the score to go above or below a certain point.

SUMMARY

To effectively study concepts and issues related to the learning of motor skills, it is important to distinguish the terms performance, which is an observable behavior, and learning, which is inferred from the observation of performance. When people learn motor skills, they typically demonstrate four performance characteristics: performance improvement over a period of time, an increase in performance consistency, a persistence of an improved performance capability for long periods of time, and the capability to adapt to a variety of performance context characteristics. The

A CLOSER LOOK

An Example of Practice Performance that Misrepresents Learning

An experiment by Winstein et al. (1996) is a good example of how practice performance may not represent the influence of a variable on the learning of a motor skill.

- **Purpose of the experiment:** Which of three different augmented feedback conditions would be best as an aid to help people learn a partial-weight bearing task? This task is a skill often taught by physical therapists. (Augmented feedback refers to performance-based information a person receives from a source external to himself or herself.)
- **The task:** The participants' goal was to learn to support 30 percent of their body weight while stepping on a floor scale with a preferred leg while on crutches. The target amount of weight was marked on the scale for each person. Participants in one group could see the scale needle move as they were stepping on the scale (concurrent augmented feedback). These participants were able to correctly adjust their weight on each trial. Two other groups received augmented feedback after performing the task (terminal augmented feedback). Participants in these groups could not see the scale needle during each trial, but saw a red line on the scale after completing one trial or a five-trial set (the five-trial group saw five red lines, each marked with the corresponding trial number of the set).
- **Practice trials and retention test:** All three groups performed eighty practice trials on one day. Two days later, they performed a retention test that consisted of twenty trials without any feedback about the amount of weight they applied to the scale.
- **Results:** *During the practice trials* the concurrent feedback group performed with very little error. The two terminal feedback groups performed with significantly more error than the concurrent group. However, *on the retention test* the concurrent group performed significantly worse than at the end of the practice trials, and worse than both of the terminal groups. The terminal feedback groups performed with about the same amount of error as they produced at the end of the practice trials.
- **Conclusion:** It is important to notice that if the retention test had not been given, the conclusion about the best augmented feedback condition for learning this task would have favored the concurrent condition. However, this conclusion would be based on performance when the various types of augmented feedback were available to the participants. The more valid way to determine which feedback condition is best for learning is when no augmented feedback is available, because it reflects the therapy goal of enabling people to perform the partial-weight bearing task in daily living conditions, which is with no augmented feedback. When the participants were tested under this condition on the retention test, the conclusion was that the *concurrent feedback was the worst learning condition.* Thus, performance during practice misrepresented the influence of the augmented feedback conditions on learning.

assessment of one or more of these characteristics forms the basis for four methods commonly used to assess skill learning. One method is to look for the improvement and consistency characteristics of performance as the person practices. We can see these when we plot performance curves of outcome and kinematic performance measures during practice. The second method is to use a retention test which assesses the persistence characteristic by requiring a person to perform a practiced skill after a period of time during which he or she has not practiced. Third, transfer tests assess a person's acquired capability to adapt to new performance conditions. A transfer test requires a learner to perform either the practiced skill in a new situation or a new variation of the practiced skill. The fourth learning assessment method involves the observation of the consistency and stability characteristics of coordination patterns during practice and on tests. This method provides an opportunity to observe previously learned and newly acquired coordination patterns as well as the transition between them.

To assess learning only on the basis of practice performance can sometimes lead to invalid inferences. Certain performance variables can artificially inflate or depress performance so that the test over- or underestimates the amount a person has learned. Additionally, a performance artifact known as a performance plateau can occur, giving the appearance that learning has stopped when it has not. Retention and transfer tests provide ways to avoid the potential problem of the influence of performance variables during practice. And, the observation of additional practice provides a means of determining if a performance plateau occurred during the learning of a skill.

RELATED READINGS

Christina, R. W. (1997). Concerns and issues in studying and assessing motor learning. *Measurement in Physical Education and Exercise Science, 1,* 19–38.

Langley, D. J. (1997). Exploring student skill learning. A case for investigating subjective experience. *Quest, 49,* 142–160.

Lee, T. D., & Swinnen, S. P. (1993). Three legacies of Bryan and Harter: Automaticity, variability, and change in skilled performance. In J. L. Starkes & F. Allard (Eds.), *Cognitive issues in motor expertise* (pp. 295–315). Amsterdam: Elsevier.

Slifkin, A. B., & Newell, K. M. (1998). Is variability in human performance a reflection of system noise? *Current Directions in Psychological Science, 7,* 170–177.

Zanone, P., & Kelso, J. A. S. (1994). The coordination dynamics of learning: Theoretical structure and experimental agenda. In S. P. Swinnen, H. Heuer, J. Masson, & P. Caesar (Eds.), *Interlimb coordination: Neural, dynamical, and cognitive constraints* (pp. 461–490). San Diego: Academic Press.

STUDY QUESTIONS

1. Explain how the terms *performance* and *learning* differ, and why we must *infer* learning from performance situations.

2. What four performance characteristics are generally present if learning of a skill has occurred?

3. What is an advantage of using transfer tests in making a valid assessment of learning? Give an example of a real-world situation that illustrates this advantage.

4. What is a performance plateau? What seems to be the most likely reason a performance plateau occurs in motor skill learning?

5. Describe a motor skill learning situation in which it may be possible to under- or overestimate the amount of learning during practice. Indicate how you would demonstrate this misrepresentation.