A more surprising finding in this research is that under some circumstances, the provision of a learning model can result in *better* observational learning than the use of an expert model. This research area was initiated by Adams (1986), who used learning models to demonstrate the performance of a manual timing task. Adams found that observation alone was insufficient for learning this task. However, considerable learning was seen if the model's KR was also presented to the observer. This is so because the observer can gain information from the model about the movement performed (both visual and auditory), from the augmented feedback presented to the model (as KR), and from seeing the success of the model's attempt to use that feedback on the next performance of the task. In this way, the observer benefits not only from "observing" the performance, but also from observing the processing operations of the model in the attempt to improve performance.

The research method used by Adams (1986) was extended by McCullagh and Caird (1990), who directly compared the effectiveness of learning models and expert models on Adams's task. Three observation groups were compared. One group had repeated exposures to a tape of a perfect execution of the timing goal. Two other groups watched a tape of a model who was learning the task; one group also received the model's KR and one did not. As illustrated in figure 11.2, the largest effects were found for those who observed the learning model and also received the model's KR (open squares). These subjects improved their performance consistently over the acquisition period, in the absence of any KR about their own performance, and both retained their performance levels and transferred to a novel timing goal better than either of the other observation groups.

These findings suggest an important application to modeling real-world tasks. Novice athletes are likely to get little insight from watching experts, other than perhaps gaining some basic information about how to perform a task. While viewing professional golf on television, for example, we get the greatest *learning* benefit from seeing these experts make mistakes. The mistakes occur so infrequently that the commentators usually replay the action and point out exactly what went wrong—what movement error resulted in the flubbed shot. In other words, the model demonstrated an incorrect action, which was accompanied by KR that identified the error. Thus, the real issue in this research may not be about the skill level of the model, but rather about what type of information is being demonstrated—errors or perfect templates of an action. It is likely that we learn more from mistakes than we do from correct performances.

Distribution of Practice

One of the variables that instructors and therapists have under their control is the scheduling of

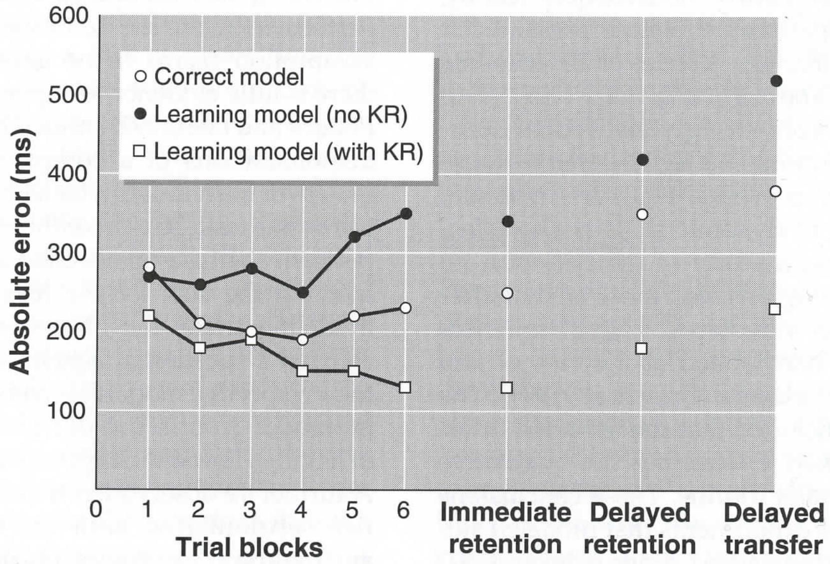

Figure 11.2. Effects of model skill level and availability of model's KR on learning.
Adapted, by permission, from McCullagh, P., and Caird, J.K., 1990, "A Comparison of Exemplary and Learning Sequence Models and the Use of Model Knowledge of Results to Increase Learning and Performance," *Journal of Human Movement Studies, 18*, 227.

periods of *work* (i.e., time spent in actual practice) and *rest* (i.e., time not practicing the task). This scheduling can be considered within a short time frame, as when one selects the amount of work and rest within a 45-min therapy session. Or the scheduling may be considered in terms of a longer scale, as when one chooses the length and frequency of sessions per week. The important question is whether or not the frequency and length of rest periods have an effect on learning the skill being practiced in the work periods. In other words, what is the best way to distribute the time spent in work versus the time spent resting—or simply, what is the best *practice distribution*?

Defining "Massed" and "Distributed" Practice

Research on practice-distribution effects has often used the terms *massed practice* and *distributed practice*. In one sense, "massing" means to put things together—in this case, running work periods very close together with either no rest at all or very brief rest intervals in between. By default, distributing practice means spacing these intervals of work apart with longer periods of rest. However, these labels are not truly satisfactory, because researchers often use these terms to describe the two extremes of practice distributions. Many experiments used more than two distribution conditions (e.g., Ammons, 1950; Bourne & Archer, 1956). Thus, these terms must be considered within the context of other conditions within any particular experiment. Different experiments, however, often established distribution conditions that were quite different from one study to another. Experiments are frequently designed such that "massed practice" involves periods of work that are substantially longer than the amount of rest between trials, eventually leading to fatigue in many tasks. For "distributed" practice, on the other hand, the amount of rest between trials often is equal to or greater than the amount of work within the trial, leading to a somewhat more "restful" practice sequence.

Virtually all the research on distribution-of-practice effects has been conducted using continuous tasks, for which the work period might be 20 or 30 s in duration. The most common apparatus for this research was the pursuit rotor tracking task. However, tasks such as mirror tracing, the Bachman ladder, and inverted-alphabet printing tasks were also popular. The effects of practice distribution using continuous tasks will be discussed first. Only a few studies have been done using discrete tasks. However, the findings are quite different from those of studies using continuous tasks, and will be presented later.

Distribution-of-Practice Effects on Performance

Many experiments were done in the 1940s and 1950s on practice-distribution effects (for a review, see Lee & Genovese, 1988). Even though these experiments involved wide differences in methods (such as the length of work and rest periods, number of trials, etc.), the results are remarkably similar. Put simply: *Given constant periods of work, short rest periods depress performance relative to longer rest periods.*

Findings from a study by Bourne and Archer (1956) are typical of the performance effects seen in experiments on practice distribution. The task was pursuit rotor tracking (see figure 2.5). Five different groups of subjects were compared; all groups had work periods of 30 s. In one group (the 0-s rest group), subjects practiced continuously for 21 trials, with no rest at all. For the other four groups, each of the work periods was interspersed with periods of rest. One group had rest periods of 15 s, and the other three groups had rest periods of 30, 45, or 60 s.

Bourne and Archer's findings were quite clear: the longer the rest period, the better the performance. Looking closely at figure 11.3, one can see that a systematic separation of the various distribution-of-practice groups had emerged quite clearly by about trial 7, and that these differences became larger with further practice. Many other examples of effects like these could be provided. Reviews by McGeoch and Irion (1952), Bilodeau and Bilodeau (1961), and Lee and Genovese (1988) describe more findings of this type.

Distribution-of-Practice Effects on Learning

For tasks such as the pursuit rotor, continuous practice would likely cause muscular fatigue to develop, and this fatigue could be expected to depress performance. In fact, looking at figure 11.3, one might argue that fatigue may have depressed performance even with some rest

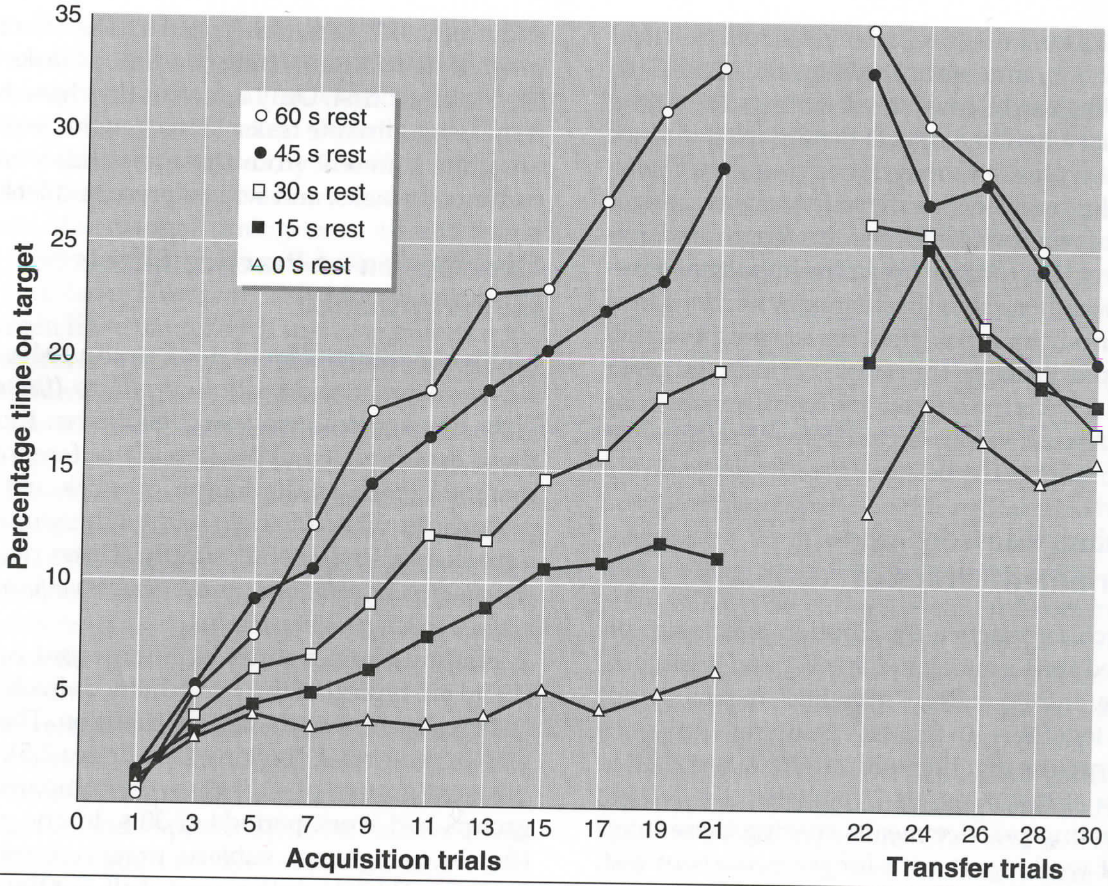

Figure 11.3. Distribution-of-practice effects on a pursuit rotor task in acquisition and retention. Trials were 30 s in duration, and separate groups received either 0, 15, 30, 45, or 60 s between practice trials. Retention trials were done with 0 s rest between trials.
Reprinted from Bourne and Archer, 1956.

between trials. Thus, because at least part of the decrement in performance displayed by these groups was due to temporary fatigue, not all of the performance depression could be attributed to differences in the relatively permanent development of skill. So, how much was due to learning?

To assess this issue, Bourne and Archer gave all of their subjects a 5-min rest period following the last acquisition trial. After this rest period, subjects performed a common transfer test in which all groups were shifted to a massed schedule—all trials were performed with 0-s rest between periods of 30 s of work. The rationale here was that if muscular fatigue was entirely responsible for the differences between groups during the acquisition trials, then the groups should be similar in performance after the dissipation of the fatigue. This was not the case, as can be seen in figure 11.3 (right-hand side).

Several items in these transfer data are noteworthy. The most important is that substantial differences were maintained between the groups after the rest period—transfer performance being increasingly better for groups that had longer periods of rest between work periods during the acquisition trials. This finding suggests that the practice distribution had a relatively permanent effect, which is supported quite well by the literature (Lee & Genovese, 1988).

Another item worth noting in these data is that the differences between the groups on the first transfer trial (trial 22) are smaller than the differences between groups on the last acquisition trial (trial 21). Thus, some of the practice-distribution effect was due to the temporary, detrimental influence of fatigue. Still, the differences due to changes brought about by learning remained large on the transfer trials. The last item to notice is that massing the transfer trials also had a depressing effect on performance. However, even after 9 transfer trials with no rest (i.e., on trial 30), the groups that had initially practiced with some rest between trials still performed better than the group that had practiced with no rest.

We have used the Bourne and Archer (1956) study to illustrate the effects of practice distribution on performance and learning. It is a particularly good example of this effect because more than two distribution groups were used and because a transfer design was used to separate the temporary from the permanent effects of the practice variable. However, several conclusions drawn from this study require further discussion (see also Lee & Genovese, 1989b).

Length of the Retention Interval

One complicating factor about the Bourne and Archer experiment is that a 5-min rest period following continuous practice may not have been long enough to allow the temporary influence of muscular fatigue on performance to dissipate (Ammons, 1988; Lintern, 1988). Thus, the transfer trials still may have been influenced by the same temporary effects that influenced acquisition performance (e.g., fatigue). A number of studies using longer rest intervals following practice, however, do not support this argument. For example, a few studies had subjects leave the lab and return later for the transfer trials (e.g., 1 day in Adams, 1952; 10 weeks in Reynolds & Bilodeau, 1952). The maintenance of distribution-of-practice effects after a period of time when these temporary effects had surely dissipated offers support for the learning difference concluded from the Bourne and Archer study.

Do the Learning Effects "Wash Out"?

The Bourne and Archer data show that performance differences in transfer begin to converge by trial 30. The convergence of effects following some transfer trials has been argued by some to cast doubt on the "relative permanency" of the learning effect. An important study by Adams and Reynolds (1954) further calls this issue into question. In this study, distributed practice was defined as 30 s of work with 30-s rest. Massed practice involved the same trial duration but with only 5-s rest. One group received 40 trials under distributed conditions. Four more groups received initial practice for 5, 10, 15, or 20 trials, respectively, under massed conditions; they then rested for 10 min, and finally transferred to the distributed-practice condition for the remainder of the 40 practice trials. Adams and Reynolds found that when the various massed-practice groups were shifted to distributed practice, they caught up (though not entirely) within a few trials to the level of performance of the group that had practiced entirely under distributed-practice conditions. A small flaw in the design, however, makes these effects difficult to interpret. The problem is that the groups that transferred to distributed-practice conditions received the benefit of a 10-min rest. The distributed group, which may have experienced some temporary fatigue effects, did not benefit from such a rest. Thus, it is difficult to know whether or not the differences that were almost "washed out" were temporary or more permanent differences.

A clever design by Ammons (1950) helps to clarify this issue. Groups received rest periods that ranged from 0 s and 20 s, up to 12 min and 24 hr between each 20-s trial on the pursuit rotor task. A 20-min rest period followed the 36th practice trial, after which subjects performed an additional 36 transfer trials with no rest between trials (many more transfer trials than had been used by Bourne and Archer, 1956). By the end of this transfer period, only small differences remained between the groups. However, Ammons (1950) asked subjects to return to the lab for *another* set of transfer trials, 1 day later. The differences that had been seen on the first transfer test—and apparently washed out by the transfer trials—were "restored" after this additional rest period. These data are a strong indicator that practice distribution has large effects on temporary performance levels and relatively permanent influences on learning.

Distributing Practice Over a Longer Time Scale

Perhaps of more direct significance to instructors and therapists are the effects of practice distribution when conducted on a much longer time scale than the single-session experiments often carried out. A few such studies have been conducted, and the results are generally similar to those of the studies done in a single session. In a very early investigation of this type, right-handed subjects were asked to throw javelins with their left hand (Murphy, 1916). All subjects practiced on 34 separate days. Massed-practice subjects performed on consecutive days (Monday to Friday) for 7 weeks. The distributed group practiced three times per week for 12 weeks. Results at the end of the 34th day of practice and on a retention test performed 3 months later showed both performance and learning benefits for the distributed group. Similar findings were reported by Baddeley and

Longman (1978) for postal workers who were training to use a keyboard. In this study, separate groups of postal workers trained for 60–80 hr using one of four schedules: work periods were conducted either once or twice per day, with the duration of each work period being 1 or 2 hr. The data for the practice period and for retention tests performed 1, 3, and 9 months later showed that the condition that massed the practice the most (2 × 2) resulted in the poorest performance and learning (see figure 11.4). Although the other three groups did not differ in these retention tests, the effects of the "most distributed" group (1 × 1) are likely diminished because practice for this group was stopped after a total accumulation of 60 hr, as compared to the 80 hr of practice for the other three groups. These data appear to suggest that there is some generalizability of the results obtained in experiments of relatively short duration to studies involving practice and retention over much longer periods of time.

Total Practice Time

From the previous sections, it would appear that it is not beneficial for learning to mass trials in the practice session. But there is another important variable that interacts with massing—the *time* involved in practice. Recall that in the experiments presented so far that used massing, the number of practice trials was held constant; and because the amount of time between practice trials was different for the massed and distributed conditions, the overall practice time was allowed to vary. That is, a group receiving massed practice will have a shorter total practice period than will an equivalent group with distributed practice.

Consider the Baddeley and Longman (1978) study just described. Although the group that practiced for 2 hr per session twice per day (2 × 2) showed the poorest acquisition and retention performance, their practice period was completed in one-half to one-quarter of the time used by the other groups to complete the training. Additional

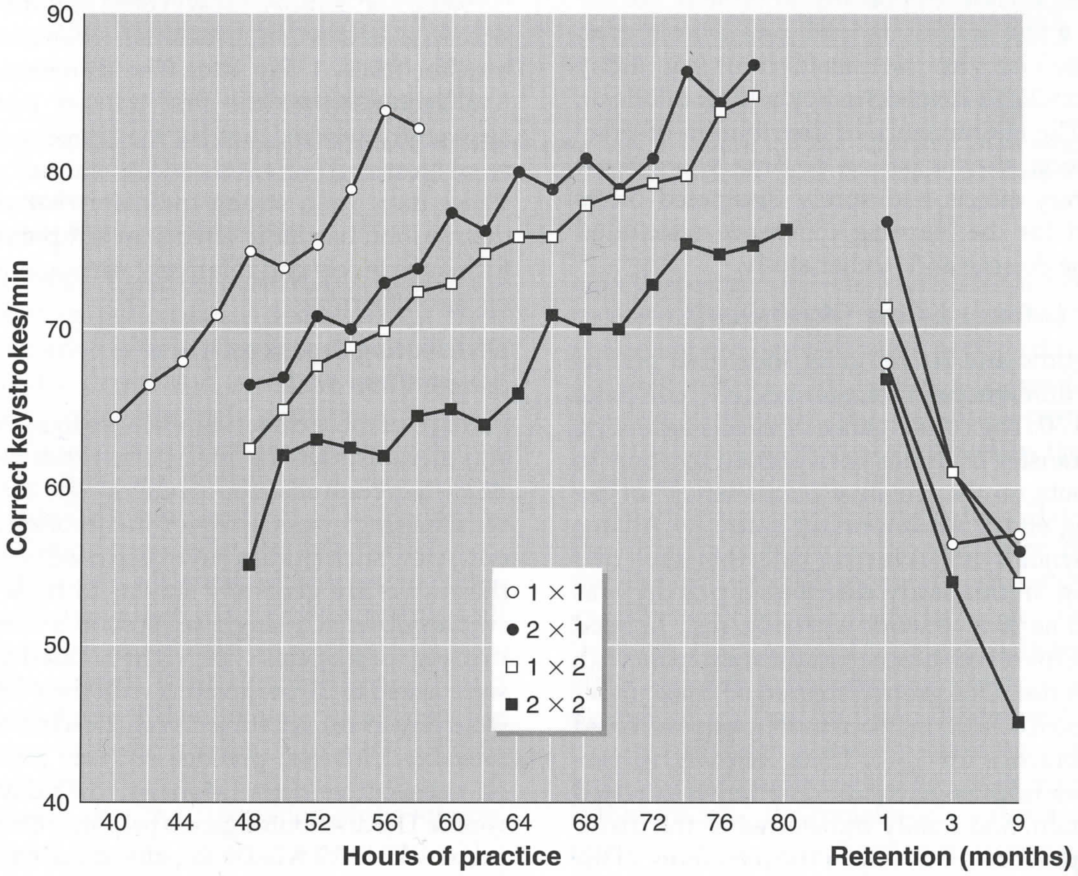

Figure 11.4. Training postal workers to use a keyboard under different practice-distribution schedules (1 × 1 refers to sessions conducted once per day of 1 hr duration each; 2 × 2 refers to two sessions per day of 2 hr duration each, etc.). Adapted from Baddeley and Longman, 1978.

training for this group would likely have resulted in improved performance and learning.

The issue of practice distribution and total practice time involves a *trade-off*. Distributed practice results in the most learning per time in training but requires the most total time to complete. Massed practice results in reduced benefits per time in training but requires the least total time.

Safety Issues

Finally, it should be clear that massing has strong effects on performance of many tasks and that the risks of injury in dangerous tasks are going to increase with massed practice. The laboratory tasks described here are not particularly dangerous, but many tasks used in sport (e.g., giant swings on the horizontal bar) and industry (e.g., work with a hydraulic paper cutter) entail considerably more opportunity for serious injury if errors are made. And most certainly for people in rehabilitation, whose motor coordination has already been affected, the risk of injury is of vital concern to the therapist. Thus, caution should be used in designing training regimes in situations in which factors such as fatigue could put the learner at risk.

Discrete Tasks

The evidence about discrete tasks is far less complete than for continuous tasks. Carron (1967, 1969) used a peg-turn task in which the subject moved 44 cm from a switch to grasp a peg in a hole, turned the peg end-for-end to reinsert it into the hole, and then returned to the key again as quickly as possible. This movement was discrete and required a movement time (MT) of from 1,300 to 1,700 ms, depending on the level of skill of the performers. Carron had subjects learn this task under two conditions: distributed (the amount of rest between trials was 5 s) and massed (the amount of rest between trials was only 300 ms, with a 5-s rest every 10 trials). Carron found no effect of the massing conditions on performance of the task on the first day while the massing was present. When he tested the subjects 48 hr later as a measure of learning, he found that the subjects in the massed condition actually performed slightly faster than the subjects in the distributed condition (1,430 vs. 1,510 ms), but it is probably more reasonable to say there were no real differences. For this discrete task, massing appeared to be neither a performance variable nor a learning variable, contrary to the rather strong effects of massing found for continuous tasks.

More recently, Lee and Genovese (1989a) investigated this apparent continuous-discrete difference directly, in parallel experiments employing very similar timing tasks. For the continuous task, there was a tendency for subjects in the distributed conditions in acquisition to perform more effectively than those in the massed conditions. This effect carried over into the delayed (7 days) transfer test, so that practice under distributed conditions in acquisition resulted in more learning, regardless of whether the transfer-test conditions were distributed or massed. This was essentially the same as had been found with the other continuous tasks (see previous section). However, for the discrete task, there was a slight tendency for the massed condition to be more effective in acquisition. Also, the learning effects on delayed transfer depended on the conditions in transfer. Even though massed practice was more effective during the acquisition phase, the distributed practice in acquisition was clearly superior for delayed transfer tests under distributed conditions. On the other hand, when the delayed transfer test was performed under massed conditions, massed practice in acquisition was clearly superior for learning. This provides at least one example indicating that massed practice can be more effective for learning than distributed practice.

It is probably premature to generalize very strongly from these two studies. But they do raise serious questions whether the effects of massed practice for discrete tasks will be simple generalizations from the wide body of findings from continuous tasks. In the first place, massing did not impair performance during acquisition in Carron's (1967, 1969) studies, and massed practice actually improved performance during acquisition in Lee and Genovese's (1989a) study. And those acquisition practice conditions that were most effective depended in a complicated way on the massing conditions in transfer. Certainly we do not understand these phenomena very well, and more studies should be concentrated on the role of these practice conditions for discrete motor learning tasks that are so highly represented in many everyday activities (e.g., kicking, throwing).

Future Research on Practice-Distribution Effects

You may have noticed that since the 1940s and 1950s when much of the research on distribution-of-practice effects was conducted, with only a few exceptions, work on this issue has stopped. Why is this the case? One possibility is that everything we could know about the topic is now known. However, this is unlikely to be true; for example, the different effects for continuous versus discrete tasks have never been satisfactorily explained. Two reasons for this lack of work relate specifically to theory testing. One reason for the decline in research in this area is that topics in learning with more exciting theoretical appeal have attracted the researchers' attention. The other reason relates to the downfall of Hull's theory (Ammons, 1947; Hull, 1943), which stimulated much of the early work in this area: Hull's theory was never replaced by another formulation that would serve as an impetus for further research (Adams, 1987; Ammons, 1988; Magill, 1988b; Newell, Antoniou, & Carlton, 1988).

It is clear that practice-distribution effects have important implications for the design of training sessions for learning motor skills. However, the applied nature of this work seems to be insufficient to drive sustained research in this area. Only when (and if) theory development resumes on this issue, it seems, will new experiments be designed and carried out.

Variability of Practice

Another factor that has been shown to affect learning is the amount of variability in a practice sequence. In one sense, this is obvious. Many tasks have variability inherent to them *(open skills)*, such as fielding ground balls in baseball or steering a car down an unfamiliar road. An important part of learning such tasks is acquiring the capability to cope with novel situations; practicing under constant (unvarying) situations would probably not be appropriate. But in another sense, this effect is not so obvious, especially when the task involves *closed skills*, for which the environmental conditions are always quite similar (e.g., archery, bowling). Here, because the criterion task to be learned is always the same, it would seem that practice under these exact conditions would be most effective for learning. Yet the evidence suggests that varied practice may be important in closed tasks as well.

Much of the research on variability of practice has been conducted to test certain predictions of *schema theory* (Schmidt, 1975b). One prediction was that transfer to novel tasks would be enhanced after practice in variable, as compared to constant, practice conditions (see chapter 13 for more on schema theory). We discuss only a few of these studies; reviews of many more of these experiments are available (Lee, Magill, & Weeks, 1985; Shapiro & Schmidt, 1982; Van Rossum, 1990).

Variability-of-Practice Effects in Retention

One way to obtain an indication of the effect of practice variability is to assess retention performance, after a period of time following the acquisition session, for one of the tasks that has been practiced. A few studies have done this by comparing the relative impacts of constant and varied practice on retention of the tasks that were practiced. There is a design complication with this type of study, however, as subjects in the different groups practice different tasks; thus what has been practiced and what is assessed in retention cannot be equated. This does not pose a problem, however, for results such as those we will see in studies conducted by Shea and Kohl (1990, 1991).

Subjects in the Shea and Kohl experiments were asked to learn to generate a goal force by squeezing a hand grip that was connected to a force transducer. In one experiment (Shea & Kohl, 1991, experiment 1), subjects performed 100 trials on the criterion task, which was to produce a force of 150 N. One group (criterion) received only these acquisition conditions. Another group (criterion + variable) received the same number of acquisition trials on the criterion task but, in addition, practiced goal forces that were ±25 or ±50 N relative to that of the criterion task (i.e., 100, 125, 175 and 200 N). Notice, however, that this variable-practice group not only had the same amount of specific practice as the criterion group, but also practiced at tasks that surrounded the criterion task—which confounds the role of the variable practice with additional practice. So, Shea and Kohl also included a third group of subjects (criterion + criterion) that practiced the criterion task, as well as performing additional practice trials on the criterion task, so that the