Motor Learning Conundrums (and Possible Solutions)

Timothy D. Lee and Laurie R. Wishart

What value is there in effective practice regimes if these same regimes do not, in themselves, engage the learner to continue to practice? This question was discussed in the context of two practice schedules that have been shown to be effective in studies of motor learning. But recent evidence has found these practice schedules to be either inefficient in their use of practice time (distributed practice), or results in metacognitive judgments of incompetence (random practice). We discuss these conundrums, possible ramifications relative to the law of practice, and potential resolutions of these issues by recent research efforts.

Although many variables influence how skill improves, no single variable is a stronger predictor of motor skill learning in adults than the amount of practice. Immediate and dramatic improvements in motor performance occur typically in the initial stages of practice, followed by smaller gains later on. That such effects are seen in most adults is rarely questioned and, in fact, have been mathematically described in terms of a power function and referred to as the Law of Practice (e.g., Crossman, 1959; Snoddy, 1926). The attainment of expertise, a level of motor prowess that is achieved by only the most highly skilled, cannot be acquired without massive amounts of practice—at least 10,000 hours of practice by some estimates (Ericsson, Krampe, & Tesch-Römer, 1993).

Many in the sporting world refer to Vijay Singh as a clear example of the beneficial effects of long hours of practice. In 2004, at age 41, Singh became the second-oldest player ever to hold the rank of world’s best golfer and the oldest golfer to become the best for the first time. This achievement is widely credited to his legendary routine of spending many hours on the practice tee, even immediately following a tournament round. When asked who the best golfer was that he had ever seen, Singh’s reply was Moe Norman, the Canadian amateur who was also legendary for the many hours of practice put in on the driving range (Potter, 2004).

But these and similar anecdotes are of little help to the instructor, coach, or therapist who work with learners who lack the opportunity to put in hundreds or
thousands of hours of practice. Some situations, by their very nature, restrict the amount of practice that can be undertaken. Two examples are a life-saving skills course that is taken over a two-day period, or an outpatient physical therapy session for an individual, post-stroke, that lasts for 45 min, three times per week for four weeks. These and many others are examples of skill learning opportunities in which the amount of total time available for practice is limited, perhaps due to instructor and/or equipment availability.

When practice time is limited, it is of primary importance that the available opportunities are used as effectively as possible, and thousands of motor learning research articles have been devoted to the study of various ways to improve motor skills. However, other issues arise when only the effectiveness of practice methods is the primary concern. Remember that practice, and the continuation of practice, is the primary factor in motor learning. How truly effective is a practice condition when, ultimately, it does little to encourage the learner to continue to practice? We address our comments in this paper to this issue. We will discuss two research issues, the distribution of practice effect and the contextual interference effect, to illustrate our main points. However, other research issues concerning practice effectiveness (e.g., augmented feedback) could also be raised.

**Distributed Practice: The Effectiveness/Efficiency Conundrum**

Perhaps no condition of practicing motor skills has received more investigation than the distribution of practice effect. The majority of these studies were conducted in the mid-20th century using tasks involving continuous movements of the upper limbs (e.g., tracking) or whole body (e.g., balance). A study by Ammons (1950) nicely illustrates the wide variety of distribution conditions that could be studied and the typical results. Participants were assigned to one of seven different experimental groups, all of which performed 36, 20-s trials of pursuit tracking in Phase 1 of the experiment. The amount of time between trials determined the different experimental conditions, with the seven groups assigned to conditions in which the rest period between each trial was 0 s, 20 s, 50 s, 2 m, 5 m, 12 m, or 24 h. The data in the left side Figure 1 illustrate the performance for these seven groups on trial 36 of practice (at the end of Phase 1). The poorest and best levels of performance were achieved by the groups with the least (0 s) and most (24 h) amounts of rest between trials, respectively. The other five groups resulted in moderate levels of success.

The relative effectiveness of these practice conditions was maintained when Phase 2 began, 20 m after the last practice trial. In Phase 2, all experimental groups now performed a series of 36 trials with no rest between trials. With the exception of the 24-h group, the beneficial effect of having had at least some rest between practice trials (in Phase 1) was mostly “washed out” by these massed (0-s rest) trials in Phase 2. This wash-out effect was only temporary, however, as can be seen in the right side of Figure 1. A retention test that was performed in Phase 3, one day later, revealed that the effects of having had distributed practice previously, during Phase 1, though seemingly “washed out” in Phase 2, had reemerged 24 hours later. The 0-s and 24-h groups again performed the worst and best, respectively, with moderate levels performed by the other distribution groups. Clearly, massed
practice was a relatively ineffective practice schedule compared to the other practice conditions.

What do these findings suggest about the efficiency of distributed practice schedules? Obviously, given the same trial durations, groups with longer intertrial intervals will require more total time to complete the schedule. The left side of Table 1 summarizes the total times required to complete Phase 1 in the Ammons (1950) study. Although the massed practice (0 s) group was found to be the least effective practice condition given an equivalent amount of practice time spent in practice, we wonder whether or not this practice schedule would have been both effective and efficient if practice had continued. Notwithstanding the effects of physical fatigue, this group could have performed 72 trials in the same total time required by the 20-s group. Perhaps a better comparison would be the relatively restful, 20-s group with the 2-m group. Over the same amount of total practice time (1 h, 24 m) the 20-s group could have performed 126 trials compared to the 36 trials performed by the 2-m group. Clearly, when relative efficiency in terms of practice time is equated, the answer about relative effectiveness is likely to take on a different understanding.

Another distribution of practice study makes the issue of effectiveness vs. efficiency even clearer. The study, conducted by Baddeley and Longman (1978), examined workers in Britain who were being trained to use a new technique for sorting mail by postal codes. The postal workers were divided into groups that
trained in sessions of either 1 or 2 h in duration, with sessions held either once or twice a day. Similar to the Ammons (1950) study discussed above, the most “massed” ($2 \times 2$) group was the least effective for both performance in training and in retention tests conducted 9 months later. Note that training in the $1 \times 1$ group was halted after 60 hours, but continued for 80 hours in the other three groups. It is unknown, but intriguing to consider, how effective the $1 \times 1$ group would have been at the end of 80 hours of practice. Nevertheless, suppose that training had ceased after 60 hours for all groups. The data in the right side of Table 1 presents the total number of days required to complete these training schedules. Once again, it is readily apparent that the relatively effective distributed conditions required either double or four times the total number of days to complete the retraining.

So what about relative effectiveness, given equal total time in practice? We have replotted some of the data from Baddeley and Longman (1978) in Figure 2, using a different metric of skill acquisition. Here, we have created three different “criteria” for the achievement of proficiency and plotted the number of days of training required by each group to reach these criteria. The data in Figure 2 illustrate that the $2 \times 2$ group achieved each criterion level of success in approximately one-half the number of days of practice compared to the $1 \times 2$ and $2 \times 1$ groups, and about one-third of the time compared to the $1 \times 1$ group. Efficiency, in terms of time to achieve the criterion, clearly favored relatively massed practice.

Performance during acquisition trials, especially so during the time when an experimental manipulation is ongoing, can be a misleading measure of learning (e.g., Schmidt, 1972; Schmidt & Lee, 1999), particularly when applied to the distribution of practice literature (Lee & Genovese, 1988). Our purpose is not to challenge or try to reinterpret this literature, but rather to raise the question that is illustrated in Table 1. If practice for the $2 \times 2$ group had continued for the same number of days as the $1 \times 2$ and $2 \times 1$ groups (for 30 days) or in the $1 \times 1$ group (for 60 days), massed practice might have been equally or more effective for both acquisition and retention.

### Table 1  Time to Complete Acquisition Period in Distribution of Practice Studies by Ammons (1950) and Baddeley & Longman (1978)

<table>
<thead>
<tr>
<th>Ammons (1950)</th>
<th>Baddeley &amp; Longman (1978)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
<td><strong>Total Time (approx)</strong></td>
</tr>
<tr>
<td>0 s</td>
<td>12 m</td>
</tr>
<tr>
<td>20 s</td>
<td>24 m</td>
</tr>
<tr>
<td>50 s</td>
<td>42 m</td>
</tr>
<tr>
<td>2 m</td>
<td>1 h, 24 m</td>
</tr>
<tr>
<td>5 m</td>
<td>3 h, 12 m</td>
</tr>
<tr>
<td>12 m</td>
<td>7 h, 24 m</td>
</tr>
<tr>
<td>24 h</td>
<td>36 days</td>
</tr>
</tbody>
</table>
There is one more important aspect of the Baddeley and Longman (1978) study that needs to be discussed. At the time of the retention test, participants in all four of the distribution groups were asked follow-up questions, such as “How satisfactory did you find your training schedule?” and “If trained again, which schedule would you choose?” For the first question, the $1 \times 1$ schedule was found to be rated as the least satisfactory and the $2 \times 2$ schedule as the most satisfactory by the participants in those groups. For the second question, the $1 \times 1$ schedule was the least popular of the alternatives that could be chosen (10%) and the $2 \times 2$ schedule was the most popular (40%). Indeed, 79% of the participants who served in the $2 \times 2$ group chose that schedule of training again, whereas only 33% of those in the $1 \times 1$ would choose to train in the same schedule again.

The open-ended nature of the follow-up questions by Baddeley and Longman makes it difficult to know exactly what it was about the $1 \times 1$ schedule that made it so unattractive to the participants, or conversely, why the $2 \times 2$ schedule was so appealing. Baddeley and Longman discussed these results as indicating something about the relative motivation that might have resulted from training in the respective schedules. If this were so, then it might be the case that a more motivating practice schedule is likely to result in continued desire to practice. If motivation to practice results in a greater likelihood that practice will continue, then according to the Law of Practice, more learning should result.

How can the effectiveness of distributed practice and the efficiency and apparent attractiveness of massed practice be resolved? We will return to this issue after discussing a different type of motor learning conundrum.
Random Practice:  
The Effectiveness/Preference Conundrum

The issue of effectiveness versus efficiency is but one conundrum faced by practitioners who might be considering the application of motor learning research in the learning of skills of daily activities. Now consider a different conundrum, regarding the benefits of blocked vs. random practice in learning new motor skills. This became a prominent research issue after Shea and Morgan (1979) published their work, although some of their findings had been anticipated by Battig (e.g., Battig, 1979). Shea and Morgan asked subjects to learn three different patterns of upper limb movements and to be able to produce these as rapidly as possible when signaled by an imperative stimulus. All three patterns were practiced for 18 trials each—the only difference between the experimental conditions (different groups of participants) was the order by which these practice trials were undertaken. In the blocked group, all 18 trials for one of the patterns were completed as a practice “block,” before the next block of 18 trials on a different pattern was practiced, and finally the last block of 18 trials was completed on the remaining pattern. The random group practiced the same three patterns as had the blocked group and for the same number of trials each. The only difference was that no more than two trials on the same pattern were practiced on consecutive trials—thus, the practice schedule required frequent switching between patterns.

The results of the acquisition trials in Shea and Morgan (1979) are presented in the left side of Figure 3. The difference between the effects of blocked and random practice orders were observed in the first grouping of trials (an average of the first three trials on each pattern)—blocked practice resulted in better performance than random practice. This blocked-practice advantage remained large throughout the acquisition trials. As such, the results were rather surprising when these subjects performed again in blocked- and randomly-ordered retention tests. These retention data are presented in the right side of Figure 3 and reveal blocked practice to be rather ineffective for retention, especially so when retention trials were ordered randomly (which probably reflects a more realistic comparison to how life skills are performed). It was clear to Shea and Morgan (1979), and to the many who have discussed, replicated, and extended these findings to other tasks since (Magill & Hall, 1990), that blocked practice was effective only for short-term gains in performance improvements. Random practice, though detrimental to immediate performance improvements, was better than blocked practice for the retention (and transfer) of the benefits of practice.

An experiment by Simon and Bjork (2001) points out the conundrum associated with random and blocked practice—concerning the impact of these practice effects on self-judgments of competence. Using a design similar to the Shea and Morgan experiment, Simon and Bjork (2001) asked subjects after every 5th trial on each pattern to be learned to make a numerical estimate regarding how well they would perform that pattern on the next day, if no further practice were to occur. These judgments of learning were conducted throughout practice and at the beginning of a retention test the following day. A summary of the actual performance results at the end of practice and in retention, together with the predicted levels of retention, are presented in Figure 4. Data presented in the left side of Figure 4 illustrates that the blocked group not only performed better than the random group,
but were also more confident than the random group of their ability to perform in retention. Similar differences in predicted levels of performance remained at the start of the retention test the next day—the blocked group was more confident in their capability to perform the retention test than the random group. However, similar to Shea and Morgan (1979), the random group actually performed the retention test much better than the blocked group. The blocked group had been grossly overconfident in their capability to perform the retention test, whereas the random group underestimated their capability.

There are a number of implications from the conundrum presented above. Previous motor learning research has clearly shown that performance during blocked practice results in an embellished level of success, relative to retention performance, and that random practice results in degraded levels of improvement, relative to retention. It is quite probable that participants in both the blocked and random groups were formulating judgments about their own learning (i.e., to perform the retention test) on the basis of current and previous levels of performance (i.e., during the acquisition trials). Similar attributions of current levels of success to predictions about motor retention have been found in other studies (e.g., Dail & Christina, 2004). And research in cognitive psychology, using a variety of paradigms, have shown that the fluency with which information is retrieved from memory is associated with predictions of recall at a future time, even though these attributions of success are often mistaken (e.g., Bjork, 1998).
What is the problem with misattributions of later success? We see at least two potential problems. First, regarding the mistaken impression that learning has proceeded better than it actually has (e.g., in blocked practice), the problematic result is a false sense of competence in the task or skill to be learned. The ramifications of this depend on the skill. A false sense of competence could result in no more than surprise or embarrassment, if, for example, the performer is trying to show off a sport skill (e.g., a figure-skating jump). However, for skills that are themselves more dangerous (e.g., “extreme” sport events) or that have serious consequences for someone else (e.g., the patient undergoing a surgical procedure), a false sense of competence resulting from previous successes in practice could be disastrous. Regarding the second problem, the mistaken impression that learning has proceeded worse than it actually has (e.g., in random practice), two potential fallouts could occur. One, steps may be taken to improve performance and thus to improve the perception that learning is being facilitated by switching, say, from random to blocked practice. Conversely, the learner may just simply quit. Either way, the danger that results lies in the cessation of practice—that more practice is no longer needed, in the case where a false sense of competence arises, or more obviously, in the case of quitting.

The issue of random practice becomes even more complicated when it is realized that people other than the learner can be involved in the decisions about how to practice. As Bjork (1998) suggests, “Doing anything during training that
increases errors or decreases the rate of improvement will not, therefore, tend to be well received—not by management, not by instructors, and not by trainees themselves” (p. 454; see also Bjork, 1994). Quite simply, poor performance during practice, regardless of whether it leads to relatively good retention or not, tends to reflect poorly on the learner and the instructor, coach, or therapist who has been responsible for setting the schedule of practice in place.

Alternatives to Distributed and Random Practice

We have suggested that the potential benefits of distributed practice are undesirable in practical settings because they are not efficient. We also suggest that random practice is undesirable to many learners because it produces metacognitive judgments that learning is not progressing well (or as well as could be achieved in other schedules). Practice schedules that are undesirable could result in reductions in the drive to practice, which further reduce the opportunity for learning. Therefore, a potential way out of these conundrums is to provide alternatives to distributed and random practice that do not reduce effectiveness and which potentially improve desirability (and motivation to practice). Some recent advances in the literature suggest that these options are worthy of continued investigation.

The rest between trials that is afforded by distributed practice sessions provides a way to take a break from performing the task, both physically and mentally. However, there is no strong theoretical rationale that a complete mental break from the task is desirable or even effective for learning. Recent studies by Shea, Wulf, and Whitacre (1999) compared the effectiveness of distributed practice trials on a balancing task with two conditions in which the periods of “rest” experienced by the distributed group were spent in observation of a cohort who was also learning the task. These two “dyad” groups either performed and watched in a block of trials (i.e., all observation trials occurred either before or after the participants’ practice trials) or alternated trials, observing one trial then practicing the next, etc. Previous research had shown that watching a cohort who was also learning a motor task for the first time was an effective learning strategy (e.g., Adams, 1986; McCullagh & Caird, 1990; McCullagh & Meyer, 1997; Pollock & Lee, 1992) and that the schedule of observation trials can also impact learning (Blandin, Proteau, & Alain, 1994; Deakin & Proteau, 2000; Shea, Wulf, & Whitacre, 1999; Weeks & Anderson, 2000). The study by Shea et al. (1999) revealed that physical practice, alternated with observation of another learner during the rest periods, benefited both acquisition performance and retention. Thus, their study showed that the efficiency and effectiveness of distributed physical practice trials could be enhanced by the introduction of observation trials during the rest periods.

Random and blocked practice represent “extreme” forms of contextual interference, and other schedules have been examined that have less extreme effects. For example, performing small blocks of trials before randomly (Al-Ameer & Toole, 1993) or serially (Landin & Hebert, 1997) rotating through the other tasks to be practiced has the advantage of facilitating learning (relative to blocked practice) but without the degrading effect on performance seen in random practice. Another method, which is contingent upon the performance of the individual, can be used to tailor the practice schedule to the learner. In the “win-shift/lose-stay” strategy, the switch to another task occurs only after the learner has achieved
a criterion level of success—the idea being that switching tasks will be most beneficial for performance and learning only after a measure of success has been attained. Although investigations using this type of strategy have only recently been undertaken, the findings are encouraging and more research using this learner-centered strategy is warranted (Simon, Cullen, & Lee, 2002).

Another strategy, which shares some relation to the win-shift/lose-stay approach, allows the learner to determine the order by which practice schedules are organized. For example, in cognitive psychology, Son (2004) found that individuals tended to mass their practice on items that were hard to remember but spaced their practice on easy items. The decision to mass or space practice trials, of course, is based on metacognitive judgments and could fall prey to the same illusions-of-competence problems as discussed previously. Nevertheless, there does appear to be some advantage for learning when the individual is given some control over their learning environment. For example, studies in which the decision about when augmented feedback should be provided is left to the discretion of the learner have shown positive influences on performance and retention, even when appropriate (yoked) control groups are included (Chen & Hendrick, 1994; Janelle, Barba, Frehlich, Tennant, & Cauraugh, 1997; Janelle, Kim, & Singer, 1995; Wrisberg & Pein, 2002; Wulf & Toole, 1999). Two studies have also investigated the use of learner-determined practice schedules in motor learning. In Titzer, Shea, and Romack (1993), a learner-determined schedule had the same beneficial effect as blocked during acquisition and was equivalent to random practice in retention, thus facilitating both performance and learning. A study by Wu and Magill (2004) revealed that the learner-determined schedule was superior to a yoked-control group, lending further support to the idea that a participant’s control over the practice schedule can have a positive influence on learning, even when the exact nature of the schedule is the same. Clearly, much more research could be done with these paradigms.

Lastly, some recent research suggests that illusions of incompetence do not necessarily need to be a byproduct of random practice. A study by Levy and Bjork (2004; see also, Ghodsian, Bjork, & Benjamin, 1997) found that interspersing transfer tests during training revealed similar findings as seen in retention tests. That is, these in-training transfer tests revealed random practice to be beneficial and blocked practice to be detrimental to learning. We see these in-training transfer tests as a potential opportunity to inform learners regarding objective indices of current competence. Perhaps using these objective measures as augmented feedback, individuals in blocked practice might better understand the false sense of competence revealed in training and for random practice, that learning is progressing better than indicated by current levels of performance.

Concluding Comments

Our goal in this paper was not to challenge the rich history of research on practice schedules in motor learning. Rather, if the measure of effectiveness combines positive retention benefits, positive metacognitive judgments, and continued motivation to practice, then it is clear that further practice schedule research is necessary. We suggest a few recent lines of research that may lead to schedules of practice that satisfy such measures of effectiveness and eagerly anticipate future research efforts.
References


