RESEARCH QUARTERLY
FOR EXERCISE AND SPORT
1988, VOL. 59, NO. 4, PP. 308-313

Massed and Distributed Practice Effects: Phenomena in Search of a Theory?

K. M. NEWELL, A. ANTONIOU, and L. G. CARLTON
University of Illinois at Urbana-Champaign

After a 25 year or so distribution of analysis interval, Lee and Genovese (L & G) have revived the massed versus distributed practice issue through a massed analysis of the empirical studies that have examined the influence of the distribution of practice conditions on the learning and performance of motor skills. As a consequence of their analysis by synthesis of the empirical literature, and buttressed by the modern technique of meta-analysis, L & G have formulated three conclusions regarding the influences of massed and distributed practice. First, distributed practice enhances performance. Second, the effects of distributed practice are larger on performance than on the first trial(s) of retention. Third, distributed practice conditions result in greater learning than massed practice conditions. The first two claims are generally well accepted in the literature leaving only the third claim regarding the benefits of distributed practice on learning as a potential contribution, although as L & G note, there have been advocates previously of a general learning effect for distributed practice. Furthermore, it should be recognized that claim three logically follows from claim two, so in effect L & G have only two independent proposals: distribution of practice facilitates learning, and distribution of practice facilitates performance of motor skills.

The conclusions of L & G rest heavily on a given interpretation of the learning-performance distinction and one's biases regarding the pros and cons of established measures of learning. The L & G analysis is also data driven. Theoretical perspectives on the massed and distributed practice effects are conspicuous by their absence. Our commentary is organized around empirical and theoretical issues that are explicit or merely implicit in the L & G analysis of the distribution of practice effects.

The L & G Synthesis of the Data

The claim that distributed practice facilitates both the learning and performance of motor skills is, at first glance, consistent with the data at hand. In particular, L & G have noted correctly that although the benefits of the distributed versus massed effect are dissipated on the retention trials, an absolute retention effect is generally still prevalent on the first few retention test trials. This interpretation assumes that absolute retention is the criterion measure for assessing learning. However, it should be noted that this learning effect of distributed practice, although reliable as reflected in the meta-analysis, is generally very small and may be regarded as trivial in both a conceptual and operational sense over the range of practice conditions that have been studied to date.

The fact that the absolute retention effects tend to dissipate within a few trials of the switch to a common distribution of practice schedule is probably what has led other scholars (e.g., Adams & Reynolds, 1954; Magill, 1985; Singer, 1980) to conclude that there is only a benefit to performance from distributed practice. Even if one wants to count this absolute retention effect as a learning effect, it should be understood that it is only a rate of learning effect. There is no evidence that massed practice produces irreversible performance losses even when transfer to distributed practice is not employed (e.g., Digman, 1959). Given the emphasis that L & G give to the permanent effects of learning as reflected in their insistence on absolute retention interval tests, it is surprising that they are willing to dismiss the transient nature of the absolute retention effects on performance during subsequent retention trials. Of course, these effects have formed the focus of what has been labeled warm-up decrement or post-rest upswing.

The conclusions drawn by L & G can be derived without resort to a meta-analysis of the data. Standard data synthesis by the researcher can achieve the same conclusion and much more, although we suspect L & G would not agree, given the opposite conclusions drawn by others on the relatively permanent effect of massed practice on learning. Indeed, as a general methodological point, we believe it is worth asking if any meta-analysis has produced a conclusion that was not anticipated (even a priori) from the standard scholarly analysis? Meta-analysis may provide a warm feeling to the researcher as a consequence of the statistical confirmation or formal backing to one's conclusions from standard data synthesis, but it should be recognized that this statistical technique can only work on the data that are provided. Meta-analysis fails to accommodate the vagaries of the extant experimental protocols which are the features that usually provide the basis for the most useful inferences from data synthesis. L & G implicitly recognize this in their follow-up discussion of some of the interesting experimental conditions that exist in the published distribution of practice literature, but their emphasis here still seems to be secondary to the conclusion of the meta-analysis, rather than a rounded assessment of independent-dependent variable relations. Meta-analysis also dictates that a number of criteria be met for a study to be included in the analysis, and as a result many of the most interesting studies may be eliminated from the analysis at hand. We now pick up on three subsidiary issues that raise further questions about the generality of the L & G conclusions regarding the effects of massed and distributed practice on the learning and performance of motor skills.

Definition of Massed and Distributed Practice. L & G provide a number of definitions of massed and distributed practice, but then without rationale, assume an empirical definition in terms of the absolute duration of the intertrial intervals in a given study to run the meta-analysis. It is our position that the very term "distributed" implies that the duration of the trial must be considered in relation to the duration of the intertrial interval. In other words, distribution of practice has to be considered on a relative, rather than absolute, basis. A 5-s intertrial interval probably has a very different effect on the performance of a task involving a single finger tap as opposed to, for example, a task requiring a continuous finger tapping performance for a duration of 5 hrs. L & G acknowledged a relative definition of massed and distributed practice, but they did not employ it due possibly to the pragmatic demands of running a meta-analysis.

If one assumes a relative definition, it is the case that in most training situations, massed practice will only be operationalized in a continuous task. In discrete tasks with short durations, such as the 500-ms movements used by Lee and Genovese (in press), it is difficult and, in most cases impractical, to impose an intertrial interval that is less than the duration of the trial itself, and, as a result, the relative massing will be small. Little is known about the distribution of practice for these tasks, and the evidence that is available suggests that the practice-rest distribution has little or no influence on the learning of short duration responses. Thus, one might argue that the categorical distinction between massed and distributed practice will, in all likelihood, be limited to continuous tasks. Even then, our observation implies that in general it is only meaningful to discuss the degree to which practice is distributed rather than inferring some absolute intertrial interval or distribution of practice to rest ratio as the critical point at which practice is either massed or distributed.

Duration of Practice Session. The majority of studies reviewed by L & G are confined to the manipulation of the duration between trials in a single practice session. Thus, the long-term effects of the distribution of practice schedules have rarely been studied, although the learning-performance distinction that L & G emphasize certainly forces the experimenter into at least a second testing session. Distribution of practice can and should also be considered with respect to the distribution of practice sessions.

Consider a study of post office workers learning a typewriter keyboard that controlled a letter-sorting machine (Baddeley & Longman, 1978). The post office workers were given 1- or 2-hr practice sessions either once or twice per day. The results revealed the benefits of distributed practice (particularly a 1-hr session once per day) in achieving any given criteria level of correct key stroke rate. Hence, this study confirmed the benefits of distributed practice sessions over a relatively long practice period (up to 3 months). These data can be used to suggest that the learning and performance benefits of distributed practice extend beyond the rather narrow time frame that was evident in most of the studies reviewed by L & G. In fact, the distribution of practice sessions in the Baddeley and Longman (1978) experiment seems to generate a significantly greater impact on learning and performance than the relatively trivial effects reported from the trial distribution studies reviewed by L & G. These effects need to be treated with caution, however, because fatigue in the later stages of a practice session may have lowered the average performance of the massed conditions.

The performance benefits of distributed practice

need to be reconsidered, however, in relation to the total duration of time that accrued from the beginning to the end of the training period. If we consider almost any criterion level of correct key stroke production in the Baddeley and Longman (1978) study, the more massed the practice the fewer number of days it took to reach the given criterion. The only exception is the 2 hrs twice per day group that did not continue practice long enough to reach the highest performance levels of the other groups. In fact, the total duration of training for the 1-hr per day group was close to 3 months, whereas the training regime of the 2-hr once per day and 1-hr twice per day was 7 1/2 weeks, while the 2-hr twice per day training regime lasted about 4 weeks.

Thus, the duration of the training period should be considered in conjunction with the actual performance level in assessing the impact of distribution of practice regimes. Performance alone, as reflected in the task criterion, is or should not always be the single criterion of significance in the training of many tasks. Indeed, the Baddeley and Longman (1978) data suggest that the interpretation of practice distribution effects on learning depends upon the criteria for learning. If learning is evaluated as a function of actual time in physical practice, distributed practice clearly leads to a greater amount learned. If, on the other hand, learning is evaluated based on performance after a number of days of practice, massed practice leads to greater learning. In a variety of contexts in which skills are learned, time is money.

The duration of the practice session is also ignored in typical manipulations of practice and rest periods within a session. Subjects practicing under a distributed practice-rest cycle will require a longer practice session to complete a given number of trials as compared to subjects performing under a massed practice schedule. Distribution of practice studies have typically kept the number of trials constant between groups while allowing the length of the practice session to vary. It is unclear what learning and performance effects would emerge if we were to choose the alternative strategy of fixing the duration of the practice session and allowing the number of practice trials to vary. This practice condition has been used infrequently in experimental studies, although it is probably the norm in the real-world training of everyday tasks. An example of a fixed session variable trial practice format is a study by Graw (1969) which had practice session length fixed at 30 min. Highly distributed practice led to lower amounts learned for both ladder climbing and stabilometer tasks. In addition, there appeared to be an optimal practice-rest ratio for each task. This study reveals how biases about how distribution of practice effects should be evaluated can lead to different interpretations regarding the influence of practice distribution on learning.

Task Specificity. L & G suggest, as others have before them (Eysenck, 1965; Schmidt, 1982), that there may be task specific influences of massed and distributed practice, and they have provided some preliminary evidence toward this point of view (Lee & Genovese, in press). A hypothesis regarding the special significance of the intertrial interval for information processing activities in discrete tasks was advanced, but both conceptual and empirical developments will need to be made if this information processing proposal is to have any theoretical impact and operational reality. Furthermore, as already indicated, task type manipulations have been confounded with the failure to adequately manipulate massed practice. This experimental problem will need to be overcome with positive findings if a task type interpretation is to hold any potential. Task type has a strong impact on the influence of a variety of learning variables on the acquisition of skill. Moreover, the study of skill learning has been dominated by a very narrow range of tasks requiring variations in only the scaling of an already established coordination function (Newell, 1985).

Eysenck (1965) has identified the task dimensions of familiarity, self-paced, and degree of drive required to perform as potential factors influencing the impact of practice conditions on performance. For example, reminiscence in pursuit-rotor learning was proposed to be due to consolidation, whereas in a task such as tapping where it was argued that no learning occurred, reminiscence was assumed to be due to the buildup and release of reactive inhibition. Unfortunately, task categorizations rest in the main on no more than intuitive assumptions and the massed versus distributed practice problem represents one of many learning issues that could benefit from a theory of tasks.

Summary. In light of the caveats briefly outlined above, we conclude that the interpretations advanced by L & G of the massed versus distributed practice effects are very strained and, to a large degree, these interpretations are contrived by convenient arguments about the definitions of learning and performance. An unfortunate consequence of the L & G paper could be the blind acceptance of the stated benefits for learning through distributed practice when empirical limitations, operational definitions, task type, and practice duration considerations suggest that a learning interpretation is premature.

The Learning-Performance Distinction

L & G have emphasized the so-called learning-performance distinction in an attempt to separate the per-

manent and temporary effects of the distribution of practice on performance. Furthermore, L & G review the various measures that have been used to reflect learning and the necessity of using transfer and retention tests to examine the permanent consequences of learning on performance. Strong conclusions are also offered regarding the independent and dependent variables that should be employed to infer the relatively permanent effects of learning. For example, they indicate that if the independent variable affects learning, then a performance effect will remain when the independent variable is no longer applied (i. e., on a retention or transfer test). Thus, for a score to reflect a true measure of learning, it must be calculated from performance data that are not contaminated by the temporary effects of the independent variable. L & G claim that only the absolute retention scores provide an adequate assessment of a true learning effect.

This strong interpretation of the learning-performance distinction seems to provide a firm basis to evaluate the massed-distributed phenomena and other learning variables such as knowledge of results (Salmoni, Schmidt, & Walter, 1984). However, this is a very old distinction predicated on a reinforcement view of learning and is not as straightforward a test of permanent and temporary effects on performance as proposed by L & G. First, statements regarding learning are always inferences from performance. This implies that veridical estimates of learning will always be problematic given that one is operating at the level of a construct, rather than an observable variable. Second, there are many attributes of a performance beyond the mere outcome scores that could be called upon to provide the basis for an inference regarding learning. Thus, the distinction between learning and performance cannot rest soley on the dimension of the task outcome no matter how convenient or traditional that measure is. Third, some definitions of learning have bypassed performance completely by focusing solely on identifiable changes to the central nervous system. In short, L & G fail to consider the many ways in which inferences regarding learning can be advanced, and, hence, the many ways in which the learning-performance distinction can be compromised.

The imposition of a rest interval of at least 1 day seems to provide, on the surface, a reasonable basis to examine the permanent influences of learning on performance. However, the rest interval introduces temporary problems of its own in the form of what has been traditionally dubbed "warm-up decrement." The first trial following the rest interval is, therefore, not a sufficient indicator of the permanent effects of learning and subsequent retention trials are contaminated by the exact practice conditions imposed on transfer. In summary the absolute retention score is not such an uncontaminated measure of learning as L & G imply leaving the learning-performance distinction as slippery as the contrast between learning and maturation (Hilgard, 1948).

Data Synthesis and Theory Development

The L & G paper is a synthesis of the empirical findings regarding massed and distributed practice. No examination is undertaken of the extant theoretical positions regarding practice schedule effects, and no attempt is made to formulate a new theoretical perspective for the massed-distributed findings. There is a well-known saying in research (we are not aware to whom it should be credited) that it takes a theory to replace a theory. That is, no matter how much the published data speak against the tenets of a given theory, the theory still stands as the guiding theoretical framework until it is replaced by another theory.

Hull's (1943) theory of learning has certainly been passed by as a significant force in learning research in general and the more limited subdomain of motor learning. However, it is clear that the Hullian concepts of reactive inhibition and conditioned reactive inhibition provided a direct theoretical perspective from which to examine massed and distributed practice (see also Ammons, 1947a, 1947b; Kimble, 1949; Rachman, 1962). What is less clear from current tests on motor learning, and certainly L & G do not help us on this issue, is whether the Hullian interpretation is ignored in the contemporary context of massed and distributed practice because of the failure of its direct predictions on these practice phenomena, or, because the theory has generally been superceded by information processing and the more recently formulated cognitive accounts of learning. Clearly, some authors have felt that the Hullian theory is still the most appropriate theoretical framework available from which to investigate the effects of massed and distributed practice (e.g., Coppage & Payne, 1981; McBride & Payne, 1979; Noble, 1978), although there were early studies claiming no learning effects for distributed practice (Adams & Reynolds, 1954), contrary to the predictions of Hull's theory.

L & G also ignore the developments by Eysenck (1965) of the Hullian concepts of reactive inhibition and conditioned inhibition into a three-factor theory of reminiscence that incorporates the additional learning construct of consolidation. This theory includes a personality dimension to accommodate individual differences to the inhibition accruing from practice. Eysenck's (1965) theoretical formulation has stimulated many

empirical studies during the last 20 years, including direct tests with Hull's theory (Ball & Payne, 1988).

It appears that the central problem with the massed versus distributed practice literature is that there is not a contemporary theory to act as a goal and a guide to ongoing empirical efforts. L & G mention the possibility of an informational account of the massed versus distributed phenomena, but there is no theoretical elaboration beyond this preliminary, almost afterthought, statement. Our inclination is that an informational account will not be sufficient to adequately approach the massed versus distributed practice phenomena. Researchers who study learning will also have to attend to a construct that they usually like to ignore by keeping it as a neutral influence—namely, motivation. Theories and models of behavior over the last 30 years have tended to account for isolated and narrow phenomena, but massed versus distributed practice is a broad-based issue for skill learning. However, it may not be broad enough to provide a useful link to the contextual interference domain as suggested in the target article. Again, L & G largely contrive a linkage without any substantive theoretical development.

An important benefit of theory is that it provides the conceptual framework from which to interpret observations of the dependent variables. L & G are adamant on the significance of the absolute retention score as a measure of learning because they view the other traditional measures of learning to be contaminated indices of learning. The validity of this interpretation of dependent variables, like any inferences beyond the data, rests on the theoretical viewpoint at hand. The difference score, for example, could be a significant dependent variable from a certain theoretical perspective. Thus, L & G's stance on the significance of the absolute retention score is based on their ongoing and publicly unknown theorizing regarding massed and distributed practice effects. To phrase it another way, there are no value-free interpretations of data, even if one is armed with the statistical technique of meta-analysis.

In summary, we can anticipate little progress on the massed and distributed practice issue unless it is attacked head on from a theoretical perspective, although it may well be that the data synthesis efforts of L & G will stimulate in the short term some further empirical efforts on, for example, the interactive influence of task type on massed versus distributed practice effects. Bottom-up, empirically driven approaches to massed and distributed practice are likely to have only temporary effects on the performance of scholars rather than the relatively permanent influences that L & G presumably seek. Indeed, Eysenck (1965) articulated a related message on this same conditions of practice issue when he said that "... the belief that good theories will materialize suddenly if only enough inductive work not guided by any theory is carried out seems to lack support in the history of science" (p. 179). The lasting conclusion we draw, therefore, from the L & G paper, albeit an implicit one, is that the massed versus distributed practice effects in motor skills reflect phenomena in search of a theory.

References

Adams, J. A., & Reynolds, B. (1954). Effect of shift in distribution of practice conditions following interpolated rest. *Journal of Experimental Psychology, 47,* 32-36.

Ammons, R. B. (1947a). Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychological Review, 54,* 263-281.

Ammons, R. B. (1947b). Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *Journal of Experimental Psychology, 37,* 393-411.

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency on training session on the rate of learning to type. *Ergonomics, 21,* 627-635.

Ball, L. J., & Payne, R. B. (1988). Comparison of inhibition and consolidation theories of spaced practice effects. *Perceptual and Motor Skills, 68,* 219-225.

Coppage, S. J., & Payne, R. B. (1981). An experimental test of current theories of psychomotor reminiscence. *Perceptual and Motor Skills, 52,* 343-352.

Digman, J. M. (1959). Growth of a motor skill as a function of distribution of practice. *Journal of Experimental Psychology, 57,* 310-316.

Eysenck, H. J. (1965). A three-factor theory of reminiscence. *British Journal of Psychology, 56,* 163-181.

Graw, H. M. A. (1969). The most efficient usage of a fixed work plus rest practice period in motor learning. (Doctoral dissertation, University of California, 1968). *Dissertation Abstracts International, 30,* 1000A-1001A.

Hilgard, E. R. (1948). *Theories of learning.* New York: Appleton-Century.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century.

Kimble, G. A. (1949). An experimental test of a two-factor theory of inhibition. *Journal of Experimental Psychology, 39,* 15-23.

Lee, T. D., & Genovese, E. D. (in press). Distribution of practice in motor skill acquisition: Different effects for discrete and continuous skills. *Research Quarterly for Exercise and Sport.*

Lee, T. D., & Genovese, E. D. (1988) Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport, 59,* 277-287.

Magill, R. A. (1985). *Motor learning: Applications and concepts* (2nd Ed.). Dubuque, IA: Brown.

McBride, D. K., & Payne, R. B. (1979). Reactive inhibition theory and intertrial correlations. *Journal of Motor Behavior, 11,* 253-259.

Newell, K. M. (1985). Coordination, control, and skill. In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory and control* (pp. 295-317). Amsterdam: North-Holland.

Noble, C. E. (1978). Age, race, and sex in the learning and performance of psychomotor skills. In R. T. Osborne, C. E. Noble, & N. Weyl (Eds.), *Human variation: The biopsychology of age, race and sex.* New York: Academic Press.

Rachman, S. (1962). Disinhibition and the reminiscence effect in a motor learning task. *British Journal of Psychology, 53,* 149-157.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95,* 355-386.

Schmidt, R. A. (1982). *Motor control and learning: A behavioral emphasis.* Champaign, IL: Human Kinetics.

Singer, R. N. (1980). *Motor learning and human performance* (3rd Ed.). New York: MacMillan.

The mailing address for K. M. Newell is: Department of Kinesiology, University of Illinois, 117 Louise Freer Hall, 906 South Goodwin Ave., Urbana, IL 61801

Research Quarterly for Exercise and Sport

Volume 59 Number 4
December 1988

1885-1985
A Century Remembered.. A Mission for the Future
American Alliance for health, physical education, recreation and dance

Research Quarterly for Exercise and Sport (ISSN 0270-1367), is published in March, June, September, and December by the American Alliance for Health, Physical Education, Recreation, and Dance, 1900 Association Drive, Reston, VA 22091. Subscriptions available to libraries, institutions, and departments at $50.00 per year. Of the membership ($60.00 per year) $20.00 covers subscription to the *Research Quarterly for Exercise and Sport.* Of the student membership ($22.00 per year) $20.00 is for subscription fees to AAHPERD Circulation Department, 1900 Association Drive, Reston, VA 22091.
Second class paid at Herndon, Virginia and additional office.
Postmaster: Send address change to *Research Quarterly for Exercise and Sport,* 1900 Association Drive, Reston, VA 22091.