ERGONOMICS, 1978, VOL. 21, NO. 8, 627–635

The Influence of Length and Frequency of Training Session on the Rate of Learning to Type

BY A. D. BADDELEY

Medical Research Council Applied Psychology Unit,
15 Chaucer Road, Cambridge CB2 2EF

and

D. J. A. LONGMAN

Post Office

Four groups of postmen were trained to type alpha-numeric code material using a conventional typewriter keyboard. Training was based on sessions lasting for one or two hours occurring once or twice per day. Learning was most efficient in the group given one session of one hour per day, and least efficient in the group trained for two 2-hour sessions. Retention was tested after one, three or nine months, and indicated a loss in speed of about 30%. Again the group trained for two daily sessions of two hours performed most poorly. It is suggested that where operationally feasible, keyboard training should be distributed over time rather than massed.

1. Introduction

It has long been claimed that short training sessions (distributed practice) produce faster learning than the equivalent amount of time spent with longer sessions (massed practice). Woodworth (1938) cites many studies which appear to support this claim across a wide range of tasks from archery to maze learning in the dancing mouse, and including the skill of learning to type. However, closer examination of the two typing studies cited by Woodworth shows that one is based on an experiment in which the groups compared differed in initial ability, making subsequent comparisons invalid, while the second turned out to be no more than a statement that this aspect of learning to type should be studied.

In recent years the tendency has been to assume that the effect of distribution of practice is primarily a laboratory phenomenon occurring only under closely specified conditions using simple laboratory tasks, and that even under these circumstances its effects may be transient. A review of the literature (Bilodeau and Bilodeau 1961, p. 263) makes this point and remarks that "An aggravation to anyone who has varied distribution of practice on a standard piece of (laboratory) hardware is the knowledge that somewhere, someone is using his findings to urge an innocent consultee to distribute the practice of his trainees as widely as possible".

As has been pointed out elsewhere (Baddeley 1976 Chapter 2) this apparent inconsistency in results of studies of distributed practice may be attributed at least in part to a failure to distinguish between two separate ways of distributing practice. The first of these is concerned with amount of practice per day, and is a variable which early studies showed to be very powerful (*e.g.* Perkins 1914). The second is the length of the interval between successive trials. Perkins (1914) found this to be a much less powerful effect. None-the-less subsequent research has concentrated almost exclusively on this variable, producing a large volume of inconclusive literature on the role of the inter-trial interval in verbal learning (*e.g.* Underwood, Ekstrand and Keppel 1964) and pursuit-rotor performance (Bilodeau and Bilodeau 1961). Over the last 20-30 y, the potentially more powerful variable of amount of training per day appears to have been almost completely neglected.

The question of optimal training schedules was raised again, however, by the British Post Office, who estimated that with the advent of mechanical letter sorting it might prove necessary to train up to 10000 operators. Since the sorting machine uses a standard typewriter keyboard, this in effect means teaching 10000 postmen to type. It is clearly important that this should be done as efficiently and economically as possible.

It seemed likely that operators would be trained in their own sorting offices, and during training might be expected to carry on with their regular jobs, rather than train full-time. With a relatively large number of men to be trained and a limited training capacity, it becomes necessary to choose between either training operators intensively a few at a time, or training a larger number of operators more gradually. The purpose of the present experiment was to provide information relevant to this decision by studying the effect of distribution of practice on learning to type.

Two lengths of training session (one or two hours) and two frequencies (1 or 2 sessions per day), were selected as being both operationally feasible and sufficiently different to show any distribution of practice effects which might occur. Thus four separate groups of subjects were trained, one group receiving one session of one hour a day (the 1 × 1 group), a second given two sessions of one hour a day (the 2 × 1 group), a third given one session of two hours a day (the 1 × 2 group), and the fourth given two sessions of two hours per day (the 2 × 2 group).

2. Procedure

2.1. *General*

For each condition, twelve postmen and six PHGs (Postman Higher Grade), were trained in groups of six. It was originally planned that all groups should train for 60 h. Since testing was on a five-day week basis this meant training periods of twelve weeks for the 1 × 1 group, six weeks for the 1 × 2 and 2 × 1 groups and three weeks for the 2 × 2 group. However, since at the end of 60 h levels of performance were not as high as had been hoped, training was extended for a further 20 h for all except the 1 × 1 group for which this proved impracticable due to previously arranged leave.

Since in the operational situation there may sometimes be a delay between completion of training and the installation of sorting equipment it is of interest to know how well this skill is retained. Subjects were therefore subsequently re-tested, either 1, 3 or 9 months after completion of training.

2.2. *Subjects*

All subjects were volunteers from the Croydon sorting office. They were postmen or PHGs who claimed neither to have used a typewriter nor to be able to play a keyboard instrument (pianists are said to be atypically adept at learning to type). Subjects were randomly assigned to the four groups, with the following contraints:

(1) Each group contained 18 subjects, 12 postmen and 6 PHGs.

(2) The groups were matched on the basis of age, the average for the four groups being as follows: 1 × 1, 35·4 y; 2 × 1, 34·1 y; 1 × 2, 34·3 y; 2 × 2, 34·6 y. Ages ranged from 19 to 46.

2.3. *Equipment*

Six portable Imperial typewriters were used. These had been modified by assigning some of the peripheral keys to specific features such as the representation of frequently used London postal districts.

2.4. *Test and training material*

This comprised the alphanumeric post code material described by Conrad and Longman (1965) together with standard typewriter training material.

2.5. *Practice and test conditions*

Once subjects had mastered the keyboard, an attempt was made to ensure that they did one test per hour. Otherwise no attempt was made to adhere to a strict timetable governing rest pauses, tests etc., for the various groups. In each case the instructor simply tried to teach the group as efficiently as possible using the standard GPO procedure for typewriter training as described in the Post Office Training Rule Book (Rg 122) H.

Since training was performed in conjunction with more normal duties, care was taken to ensure that time of training session (a.m. or p.m.), and shift worked (early, late or regular day) were balanced across conditions as far as was possible with the different training schedules. The two sessions in the 2×1 h and 2×2 h groups were always separated by at least 2 h, one being in the morning and the other in the afternoon. One subject withdrew from the 1×1 h group, and data from a second 1×1 subject had to be discarded since he was without his reading glasses for a month and hence could not read the test material adequately.

3. Results

3.1. *General*

The simplest way of expressing the results would be to present a graph showing for each group the mean performance on each hour, averaged over all operators. In principle this will be done. There is however, the problem of days on which operators were, for some reason, absent *e.g.* because of a public holiday, or an illness.

The first procedure adopted has been to treat successive working days as a continuous sequence. Thus a man in the 1×1 group for example, who was never absent, would yield performance data for 60 h. A man who had two days absence would have a record which terminated after 58 h. This procedure does give a realistic indication of rate of learning, but has the disadvantage that the last few hours may have so few men represented that the data are of little value. Furthermore, it might be argued that from a practical point of view the possibility that one schedule was particularly liable to be affected by days missed due to absence, should be taken into account in assessing the value of the schedule. For this reason all statistical comparisons are based on a nominal 60 or 80 h of training, regardless of how many of these training hours were in fact lost. Thus the graphs give an indication of rate of learning per *actual* hour of training, while the tables and statistical tests show the rate of learning per *nominal* hour, ignoring the fact that most subjects lost one or two hours during the training programme.

3.2 *Performance after 60 hours*

3.2.1. *Stage 1.* Stage 1 refers to that period of training prior to the use of lists containing all codes. During this period operators could not be said to be 'operational' since there were still procedures and parts of the keyboard they had not yet adequately learned. Progress during this phase is basically governed by the instructor who decides at what point new material should be introduced. For administrative convenience it is almost essential that a class advances together and it is up to the instructor to decide at what moment to move on.

Table 1 shows the mean number of actual hours of training required to learn the whole keyboard under the four experimental conditions, together with the range of values obtained.

Table 1. Mean number of hours needed to learn the keyboard as a function of training schedule

	Schedule			
	1 × 1 h	2 × 1 h	1 × 2 h	2 × 2 h
Mean hours to learn keyboard	34·9	42·6	43·2	49·7
Range	26–44	34–46	37–45	46–54

The most striking feature of this result is the slowness of the 2 × 2 h schedule. The fastest subject in this group is slower than the slowest in the 1 × 2 or 1 × 1, which is easily the fastest group.

3.2.2. *Speed.* Average speed of performance on test runs for the four groups is shown in Figure 1, together with the equivalent data from the typewriter keyboard group in the Conrad and Longman (1965) study. Since observations based on only a few subjects can be very misleading, points with data from less than six operators have been omitted. As with Stage 1, the most striking feature is the slowness of the 2 × 2 group, especially compared with the 1 × 1 group.

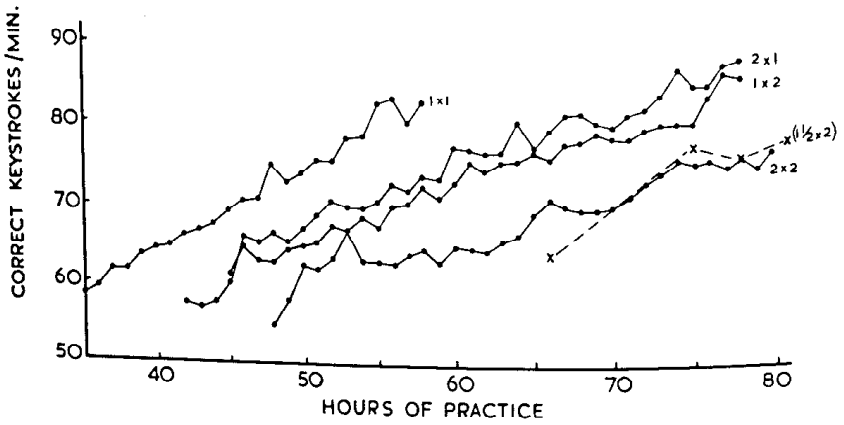

Figure 1. Rate of acquisition of typing skill as a function of training schedule. The 1·5 × 2 function is based on Conrad and Longman's (1965) typewriter group who trained for two sessions of 1·5 h per day.

A statistical analysis was performed using each subject's speed on the test following a nominal 60 h of training. The mean speed for the four conditions together with the mean number of actual hours of training are given in Table 2.

Table 2. Mean Rate (Correct Keystrokes/min) after a nominal 60 h of training

	Schedule			
	1 × 1 h	2 × 1 h	1 × 2 h	2 × 2 h
Mean Rate	79·31	73·43	71·12	64·78
Range	53·3–95·7	54·4–98·1	43·9–101·2	46·6–82·3
Mean actual hours	55·2	54·3	53·5	58·9
Range	45–60	46–58	46–58	50–60

Analysis of Variance showed a highly significant effect of length of session, with sessions of one hour being better than sessions of two ($p < 0{\cdot}01$), and a smaller effect of

frequency of session, with one per day being better than two ($p=0{\cdot}05$). There was no significant interaction. Comparison between groups, using a t test, showed the 1×1 h group to be faster after 60 h training than the 2×2 h group ($p<0{\cdot}01$) or the 1×2 h group ($p<0{\cdot}05$, 1 tail), and the 2×1 h group to be reliably faster than the 2×2 h group ($p<0{\cdot}05$).

3.2.3. *Errors.* (*i*) Uncorrected Errors. From a practical point of view these are the most important type of error, since they may lead to a mis-sort if the error produces a permissible post code. Their occurrence in the four groups is shown in Figure 2. Results after 60 h for the four groups are shown in Table 3. Since error scores are based on relatively small numbers of error responses, data from the last five tests were used so as to give a more reliable measure.

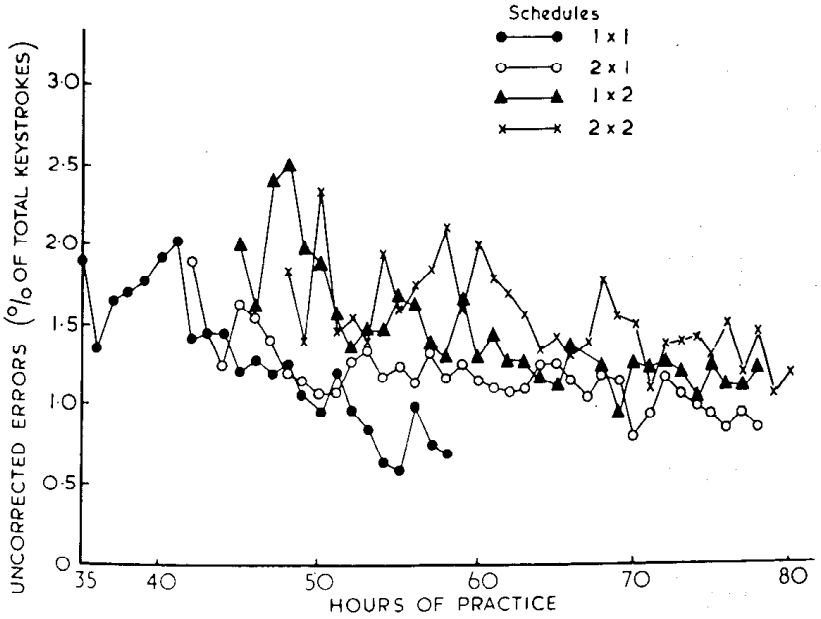

Figure 2. Mean rate of uncorrected errors as a function of training schedule.

Analysis of Variance showed length of session, frequency of session and the interaction between length and frequency all to be highly significant ($p<0{\cdot}001$). This is almost certainly due to the very high error rate found in the 2×2 h group, which produced significantly more uncorrected errors than any of the other groups ($p<0{\cdot}01$ in each case). No other inter-group difference was significant.

Table 3. Mean Percent Uncorrected Errors after a Nominal 60 h of Training

	Schedule			
	1 × 1 h	2 × 1 h	1 × 2 h	2 × 2 h
Percent uncorrected error	1·09	1·14	1·41	2·06
Range	0·22–2.18	0·06–2.42	0·40–3·45	0·38–4·65

(*ii*) *Corrected Errors.* In all conditions subjects made errors which were immediately detected and corrected. Such errors constituted about 0·5% of keystrokes in all groups, this level remaining remarkably stable throughout training. This result is consistent with Rabbitt and Vyas's (1970) suggestion that subjects use detected errors as a means of monitoring performance.

3.4. *Performance after 80 hours*

Since all but the 1 × 1 group continued up to a nominal 80 h of training, a second analysis was performed at this point. Table 4 shows the mean keying rate at this stage, excluding four subjects who had missed 20 or more hours of training. Statistical analysis confirms the results of tests after 60 h in showing a significant difference between the 2 × 1 and the 2 × 2 groups ($p < 0{\cdot}05$) but not between the 1 × 2 and the 2 × 2 groups.

Table 4. Mean Rate (Keystrokes min^{-1} after a nominal 80 h of training.

	Schedule		
	2 × 1 h	1 × 2 h	2 × 2 h
Mean Rate	89·4	82·8	77·6
Range	68·5–119·8	47·8–117·5	55·4–93·2
Mean Hours Actual Training	77·4	75·4	79·2
Range	64–80	69–80	70–80
No. of Subjects	16	17	17

3.5. *Retention*

Subjects from each group were divided into three sub-groups of approximately equal skill as measured by performance at the end of training. One such sub-group was retested after approximately 1 month, one after 3 months and the third after 9 months.

Retesting involved a session of one hour in which subjects first spent a few minutes 'warming-up' by practising simple phrases etc., followed by a 15 min timed test run on post code material. They were then given a number of short stimuli which aimed to

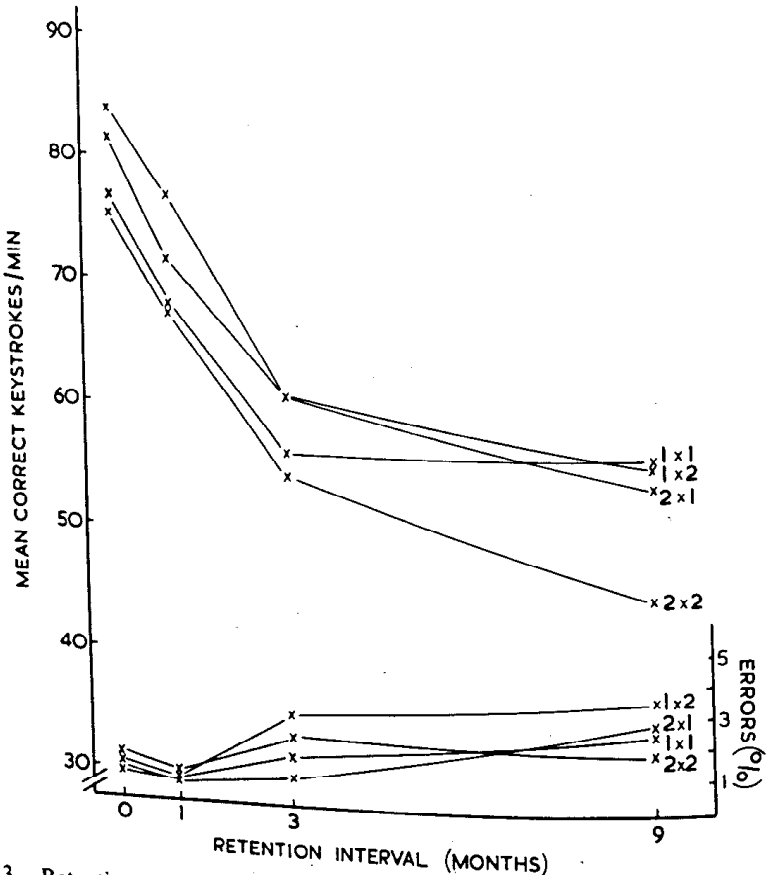

Figure 3. Retention of typing skill over a 9 month period as a function of training schedule.

remedy any obvious defects shown on the first run. In the last 15 min they performed a second test run. Figure 3 shows performance on this second run, which seems likely to give a truer indication of their retained skill than the earlier run. It proved possible to retest 62 of the 70 subjects who completed training. Of these 7 claimed to have used a typewriter since completing training and their retention data were therefore discarded.

It is clear that some forgetting does occur, but the results seem to suggest that the loss of speed is not excessive (approximately 30% after 9 months), and does not increase substantially after the first three months (with the possible exception of the 2×2 h group). Similarly, although error rate does increase, for most subjects it remains within reasonable limits.

3.6. *Subjective Ratings.*

At the time of retention testing, subjects were asked to rate the schedule on which they had trained as '*very satisfactory*,' '*satisfactory*', '*reasonable*', *unsatisfactory*' or '*very unsatisfactory*'. The first section of Table 5 shows the results of this rating. While subjects in general responded favourably, the trend is clearly in the opposite direction to that found in terms of learning efficiency, with the 1×1 h per day group producing a less enthusiastic response than the 2×2 h per day group. This is also reflected in the second question regarding the preferred schedule (see Table 5): subjects tended to prefer the schedule they had experienced, but this is much more pronounced in the 2×2 h group than for the 1×1 h schedule. A similar pattern is discernable on the third

Table 5(a). Responses to the question
"How satisfactory did you find your training schedule?"

Schedule	Very Satisfactory 1	2	Reasonable 3	4	Very Unsatisfactory 5	Mean
A. 1×1 h/day	7	1	3	2	2	2·40
B. 2×1 h/day	6	6	0	2	0	1·86
C. 1×2 h/day	7	4	3	2	0	2·00
D. 2×2 h/day	7	5	3	0	0	1·73

Table 5(b). Responses to the question
"If trained again, which schedule would you choose?"

Schedule Trained	Schedule Chosen A	B	C	D
A. 1×1 h/day	5	1	6	3
B. 2×1 h/day	1	6	3	4
C. 1×2 h/day	0	2	8	6
D. 2×2 h/day	0	3	1	11

Table 5(c). Responses to the question
"How keen would you be to undergo further training on the *same* schedule?"

	Be very keen 1	2	Take part if necessary 3	4	Refuse 5	Mean
A. 1×1 h/day	9	4	0	0	2	1·80
B. 2×1 h/day	9	2	2	0	0	1·46
C. 1×2 h/day	11	5	0	0	0	1·31
D. 2×2 h/day	11	3	0	0	0	1·21

Table 5. Subjective ratings on the four schedules studied. Numbers represent the distribution of responses to the three questions as a function of training schedule. Not all subjects could be contracted, and some subjects failed to respond to one or more questions.

question, regarding willingness to undergo further training. In general, enthusiasm seems to be high with the exception of two 1 × 1 h subjects who would refuse further training.

Two conclusions may be drawn from this pattern of results. First, is seems unlikely that the better performance of the 1 × 1 h group is due to a higher level of motivation. If anything they were less motivated and resented having to spend 12 weeks learning what some of their colleagues were learning in 3 weeks. This should obviously be borne in mind in evaluating possible schedules, although it should not be given too much weight since motivation remained high even in this group, and might have been even higher if subjects had not had colleagues simultaneously undergoing a more rapid training. Secondly it is worth noting that an evaluation of the various schedules purely in terms of subjective ratings would have led to a recommendation of a schedule which is associated with the slowest learning, the least accurate performance and the poorest retention. As such it amply reinforced Poulton's (1976) warning of the danger of relying exclusively on subjective assessments.

4. Discussion

The decision as to what schedule should be used to train operators will clearly depend heavily on operational considerations. Nevertheless, the results of the present study are sufficiently clear-cut to allow certain recommendations. In general, a session of 2 h seems to be too long for efficient training. One hour appears to be a more satisfactory length. One session per day is slightly more effective than two, though this is not a very great difference. It is clear that the 2 × 2 h training schedule is not a good one. Despite the fact that fewer hours were lost on this schedule due to absences, holidays etc., it produced consistently poor performance in both speed and accuracy (the 1 × 1 group was faster and more accurate after 60 h training than the 2 × 2 group after 80 h). It is interesting to note that the typewriter keyboard group in the Conrad and Longman (1965) study shows a very similar rate of progress. This group was given two sessions of 1·5 h per day, separated by a half-hour break.

The retention tests indicate that keyboard training is reasonably well retained. Although both speed and accuracy deteriorate if the task is not practised, the loss is not enormous.

In general terms, this study supports those carried out in the early years of this century in suggesting that it is a mistake to try to cram too much training into a single session, or indeed a single day. At a more theoretical level it suggests that preoccupation with the small and unreliable inter-trial interval effect has distracted attention from the more robust and powerful effect of the limit on amount of learning that can be accomplished per day. Such effects are likely to be of considerable practical importance; they raise many further questions, including: (1) If two tasks are to be learned concurrently, how similar must the second task be to impair the rate of learning of the first? (2) Is the effect limited to discrete motor skills such as typing? (3) How much of the disadvantage of massing can be accounted for by intra-day changes in instructor (or instruction) effectiveness? Since there is at present no theory of learning which predicts such effects they also represent a considerable theoretical challenge.

The whole of this 13 month study was carried out at the Croydon sorting office and was made possible by the generous co-operation of everyone concerned. We are particularly grateful to the Head Postmaster, Mr. E. Caddy, OBE., TD., to the Assistant Head Postmaster, Mr. W. H. Line, to the Chief Inspector, Mr. H. H. Scott and his colleagues, to the local representatives of the UPW who gave both general support and practical help, and of course to the men who took part.

Quatre groupes d'employés des Postes ont été entraînés à taper des données en code alpha-numérique à l'aide d'un clavier de machine à écrire traditionnel. Cet entraînement a été effectué au cours de séances d'une durée de une ou deux heures ayant lieu une ou deux fois par jour. L'apprentissage a été le plus efficace pour le groupe bénéficiant d'une seule séance d'une heure par jour et le moins efficace pour le groupe travaillant en deux séances de deux heures par jour. La rétention a été testée après un, trois ou 9 mois. La perte de la rapidité était d'environ 30%. Là encore le groupe 'deux séances de deux heures' s'est avéré être le moins performant. On suggère qu'il vaut mieux instaurer un apprentissage distribué dans le temps, plutôt que massé, du moins en ce qui concerne l'apprentissage de frappe sur un clavier.

Vier Gruppen von Postangestellten wurden in der Eingabe eines alphanumerischen Codes in eine konventionelle Schreibmaschinentastatur eingeübt. Die Einübung erfolgte in ein- bis zweimal täglichen Perioden von ein bis zwei Stunden Dauer. Die Übung war besser in der Gruppe, die eine einstündige Trainingsperiode pro Tag hatte, und schlechter für die Gruppe, die zweimal zweistündige Trainingsperioden absolviert. Die verbliebene Übung wurde nach einem, drei und neun Monaten ermittelt und zeigte einen Verlust an Arbeitsgeschwindigkeit von ungefähr 30%. Auch hier wiederum arbeitete die zweimal täglich je zwei Stunden eingesetzte Versuchsgruppe schlechter. Wo immer es durchführbar ist, wird für ein Maschinentastaturtraining das verteilte Üben gegenüber dem massierten Üben empfohlen.

References

BADDELEY, A. D., 1976. *The Psychology of Memory*. New York BASIC BOOKS.

BILODEAU, E. A., and BILODEAU, I. M., 1961. Motor Skills Learning. *Annual Review of Psychology*, **12,** 243–280

CONRAD, R., and LONGMAN, D. J. A., 1965. Standard typewriter versus chord keyboard—an experimental comparison. *Ergonomics*, **8,** 77–78.

PERKINS, N. L. 1914. The value of distributed repetitions in rote learning. *British Journal of Psychology*, **7,** 253–261.

POULTON, E. C., 1976. Quantitative subjective assessments are almost always biased, sometimes completely misleading. *Bulletin of the British Psychological Society*, **29,** 385–387.

RABBITT, P. M. A., and VYAS, S. M. 1970. An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. In A. F. Sanders (Ed.), *Attention and Performance 111,* 56–76. Amsterdam: NORTH HOLLAND.

UNDERWOOD, B. J., EKSTRAND, B. R., and KEPPEL, G., 1964. Studies in distributed practice: XXIII. Variations in response-term interference. *Journal of Experimental Psychology*, **68,** 201–212.

WOODWORTH, R. S., 1938 *Experimental Psychology*. (London: METHUEN).

Manuscript received 1 February 1977.
Revised manuscript received 11 July 1977.