Feedback-Induced Variability and the Learning of Generalized Motor Programs

Gabriele Wulf & Richard A. Schmidt

Max-Planck Institute for Psychological Research, Munich, Germany
Department of Psychology, University of California, Los Angeles, USA

Available online: 14 Jul 2010

To cite this article: Gabriele Wulf & Richard A. Schmidt (1994): Feedback-Induced Variability and the Learning of Generalized Motor Programs, Journal of Motor Behavior, 26:4, 348-361

To link to this article: http://dx.doi.org/10.1080/00222895.1994.9941691

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Feedback-Induced Variability and the Learning of Generalized Motor Programs

Gabriele Wulf
Max-Planck Institute for Psychological Research
Munich, Germany

Richard A. Schmidt
Department of Psychology
University of California
Los Angeles

ABSTRACT. As compared with providing extrinsic feedback on each of a set of practice trials, reducing the feedback frequency in various ways facilitates long-term retention. One explanation is that frequent feedback operates proactively on the subsequent trial, inducing excessive variability that degrades learning. We tested this view by giving or not giving feedback "reminders," where the postresponse feedback was given again just before the next attempt at a task. The reminder manipulation was examined in both blocked and randomized practice sequences during the learning of three limb-patterning tasks. Reminder feedback increased response variability during practice in both random and blocked practice. As measured in retention tests, though, feedback reminders degraded learning in random practice, but not in blocked practice. This implies that the frequently found learning advantages of random, as compared with blocked, practice might be due to feedback's facilitation of retrieval operations in blocked practice (as well as in random practice with feedback reminders); such overfacilitation of retrieval operations has been shown to degrade learning. Additional analyses revealed that random, as compared with blocked, practice enhanced the learning of the fundamental pattern of action (generalized motor program) but had little effect on the ability to scale the pattern in amplitude or time (parameterization).

Key words: contextual interference, feedback, generalized motor programs, motor learning, variability

Over the past few years, many researchers have examined various scheduling manipulations of experimenter-supplied (i.e., extrinsic) feedback during the learning of laboratory motor tasks (see Schmidt, 1991a, for a review). The usual comparison condition has been one with feedback presented after every trial—a condition that, until the mid-1980s, had been considered optimal for learning (e.g., Adams, 1971; Bilodeau, 1966). For example, feedback has been reduced in frequency; that is, it has simply been withheld on selected trials during practice (Winstein, 1988; Winstein & Schmidt, 1990; Wulf & Schmidt, 1989). Also, feedback has been presented in summary form (after Lavery, 1962) in experiments where the learner has had to wait for a set of several trials before being informed about the proficiency of each of them (Schmidt, Lange, & Young, 1990; Schmidt, Young, Swinnen, & Shapiro, 1989). In a variant of this summary technique, the average score for the set of responses was presented only after the set had been completed (Young, 1988; Young & Schmidt, 1992). The important finding for the present purposes is that, as compared with every-trial feedback in practice, all of these conditions produced more effective learning as measured on retention or transfer tests, particularly so if the test was delayed by a few days.

These findings run strongly counter to the common belief that emerged from feedback research through the 1970s. Under this earlier perspective, nearly any factor that made the feedback more immediate, more frequent, more informationally rich, or in some sense more useful for correcting errors, was supposed to enhance learning. Such feedback variations usually did, in fact, enhance performance during the acquisition phase when feedback was being manipulated between groups (see Bilodeau, 1966, or Newell, 1976, for reviews). But, as Salmoni, Schmidt, and Walter (1984) later pointed out, examining the effects during acquisition confounds the relatively permanent effects of feedback (due to learning) with any...
transient effects of feedback that could vanish on a retention test or when the feedback is later withdrawn. However, examining performance on delayed retention tests with feedback conditions equated between groups separates the relatively permanent effects from transitory ones, allowing any differences between groups at retention to be interpreted as differences in the relative amount learned during acquisition. Evaluating learning on retention or transfer tests revealed that increased feedback frequency and “usefulness” in fact degraded learning rather than facilitated it. More importantly, these new findings also raised important questions about the nature of underlying processes allowing less frequent and “useful” feedback to benefit learning, which is the general focus of the present article.

Several interpretations of these frequent-feedback effects are possible (Schmidt, 1991a). Some of these emphasize what can be termed retroactive memory processes—a way of thinking about feedback and learning that has dominated thought in this area for decades. Such views focus on the relationship, or connection, between the feedback and the action that produced it. In a sense, the feedback is thought to “work backward” in relation to the action to which it belongs, encouraging researchers to focus on processes occurring between the action and the feedback.

One version of this notion is that frequent feedback distracts, or even blocks completely, certain postresponse information processing activities that would occur if feedback were not provided on a particular trial. For example, frequent feedback might interfere with the subjective analysis of intrinsic (i.e., response-produced) feedback, preventing the development of error-detection processes (see Schmidt et al., 1989, for evidence). Such self-correcting processes would seem important for performance in retention, particularly if extrinsic feedback is withdrawn. Related is the finding that providing feedback instantaneously after an action was less effective for learning than delaying it by a few seconds (Swinnen, Schmidt, Nicholson, & Shapiro, 1990); here, the instantaneous feedback presumably blocked processing of intrinsic feedback, whereas a short delay allowed some of this processing to occur. A different retroactive view is that summary or average feedback filters the error information, allowing the feedback to be a better descriptor of the subjects’ fundamental action pattern than would error on any one trial.

A second general way to conceptualize feedback utilization involves what can be termed proactive memory processes. These newer ideas focus on the relationship between feedback and the next action, where feedback is viewed as “working forward” in time to produce its effects. Such a view does not exclude the retroactive processes discussed above, as both kinds of activities could operate during practice, but at different times. One proactive hypothesis is that frequent feedback might provide too much facilitation of next-response planning and retrieval—acting almost as a hint—thereby reducing the subject’s need to perform memory retrieval operations (retention practice; see Bjork, 1975) thought to be critical for learning. If reduced feedback enhances the subject’s tendency to produce retrieval operations during practice, retention performance would be facilitated relative to an every-trial feedback condition.

A second proactive hypothesis—the one we focus on most strongly in the present article—is that frequent feedback produces excessive variability during practice. This view comes from the old notion that providing feedback induces change whereas withholding feedback induces repetition (Bilodeau, 1966). Additional evidence comes from Nicholson’s (1992; Nicholson & Schmidt, 1991) findings that feedback after an action induces instability (see also Bilodeau, Sulzer, & Levy, 1962, for initial hints at such effects). Nicholson showed that, if the learner is instructed to repeat an action (even if the action was in error), providing (versus not providing) feedback about the just-previous response increased variability in the repetition by nearly 50%. Remarkably, these effects occurred in several variations of the paradigm, including a case where the learner was informed that the upcoming trial was to be repeated and that any feedback about it was to be ignored.

If feedback produces variability, then frequent feedback could be thought of as producing excessive modifications from trial to trial, preventing the learner from generating stable behavior. These trial-to-trial changes driven by feedback may be beneficial in the short term to keep the behavior near the target (preventing drifts off target or repetitions of over- or undershoots, for example), but they could induce instability that interferes with learning. In addition, the feedback report can be conceptualized as containing a “true”-score component plus a random error based at least in part on variability in the low-level neuromuscular processes that generate muscle forces and timing. Such variability is probably largely uncontrollable and seems largely insensitive to practice (Schmidt, Zelaznik, Hawkins, Frank, & Quinn, 1979). But the true-score component does change with practice, so that the random component later in practice might make up a large proportion of the total error in the feedback report. Even so, to the learner, feedback represents an error to be corrected, so that frequent feedback could produce many changes driven by small, largely uncontrollable random processes. In this way, frequent feedback may be maladaptive in the long term because these trial-to-trial variations interfere with the development of stable memory representations and the capability to perform well at the retention test.

The experiment reported here evaluates this proactive role of frequent feedback by examining feedback variations in practice that should manipulate response variability. We manipulated response variability in acquisition by giving or not giving so-called reminder feedback (after Bilodeau & Bilodeau, 1958), where the feedback...
from a previous trial was given again—as a reminder—just before the next attempt at the same action. This extra feedback, particularly in acquisition where several different tasks are presented in a randomized order, might induce additional variability over and above a no-reminder condition. If so, the manipulation of reminder feedback during acquisition should provide a way to examine the hypothesis that feedback-induced variability degrades learning, where learning would be evaluated later on a retention test.

This procedure provides the interesting prediction that giving reminder feedback should degrade learning in randomized practice schedules. This is certainly contrary to earlier views of the role of “useful” feedback for learning, where it would have been expected that feedback reminders would have been beneficial. (Indeed, Bilodeau & Bilodeau [1958] originally used reminders as a control for the supposedly detrimental effects of feedback forgetting.) In blocked practice, however, numerous trials of the same task are presented successively, and reminder feedback should be largely redundant with the usual feedback given after each trial because the feedback and the reminder provide the same information separated only by a few, uninterrupted seconds. Here, reminder feedback should not induce as much variability, and, compared with the no-reminder condition, it would not be expected to degrade learning. Thus, the experiment involved the examination of the effects of induced variability as created by reminder feedback on practice schedules that were either randomized (different tasks on successive trials) or blocked (same tasks on successive trials). The effect of reminder feedback would most likely act forward in time, operating on the trial immediately following the reminder, allowing increased insight into the proactive processes in feedback's operation.

A final goal of this article is the evaluation of the nature of the practice effects in terms of the theory of generalized motor programs (GMPs; e.g., Schmidt, 1975, 1988). Gains in skilled action can be caused either (a) by the enhanced accuracy of a fundamental movement representation (the GMP) or (b) by the increased effectiveness with which the GMP is parameterized at the time of test. Parameterization refers to relatively superficial features, such as the overall speed or amplitude of a given pattern. Of course, both effects could be produced at the same time, and are usually confounded. In this article, we describe a method of separating these effects that allows us to examine the role of practice scheduling and reminder feedback on the learning of GMPs and the capability to parameterize them.

Method

Practice variability was induced in acquisition by giving or not giving reminder feedback during acquisition. This manipulation was crossed with blocked versus random practice. The reminder feedback should induce response variability for the random conditions, as it is a repetition of information presented several trials earlier. In blocked conditions, however, reminder feedback should be far less effective in inducing variability because it is simply a repetition of information presented a few seconds earlier. If feedback-induced variability operates proactively to degrade learning, reminder feedback should impair learning for random practice, but not for blocked practice, with learning measured on delayed retention and transfer tests. New methods allowed an analysis of the extent to which any treatment effects were due to GMP learning or parameterization learning.

Subjects

Subjects were 56 students from the University of Munich who were paid DM 15.00 for their services. Subjects were not informed about the purpose of the experiment and had received no prior experience with the experimental apparatus.

Apparatus

The apparatus consisted of a wooden lever attached at one end to an aluminum base that was mounted on an almost frictionless vertical axle turning in ball-bearing supports. The supports were mounted on a table, allowing the axe to move freely in the horizontal plane over the table surface. A vertical handle was attached to the other end of the lever. The position of the handle could be adjusted so that the subject’s forearm rested on the lever, with the elbow aligned over the axis of rotation. A potentiometer was attached to the lower end of the axe to signal lever position; its output was fed to a Hewlett Packard Vectra QS/20 computer, where it was sampled at 200 Hz. A cover was placed over the apparatus to prevent subjects from seeing the lever while they were seated at the table.

Task

Subjects were asked to learn predetermined goal lever movement patterns, produced by movement at the right elbow. There were three different movement patterns, as shown in Figure 1. The goal movement time for all three patterns was 937 ms, and all movements started and ended at a lever position of 39°, where 0° was parallel to the subject’s frontal plane. Otherwise, the spatiotemporal requirements were quite different for the three patterns. That is, they can be assumed to have required different GMPs (Schmidt, 1988), in that there exists no linear scaling factor (or parameter) that would transform any one pattern into any of the others. The goal movement patterns (termed templates) and the feedback were displayed on a computer screen (EIZO Flexscan 9060S).

Procedures

Before each trial, the template of the next task to be attempted was displayed on the screen in white (against a blue background); it appeared suddenly and completely, and was not drawn in real time. The template was
removed after 4 s, replaced by two 2.5-cm-long vertical lines; one fixed line represented the starting position, and the second line, movable by the lever, represented the lever's actual position. The subject was then to move the lever so that the two lines were aligned at the starting position. Once alignment was achieved, a tone was presented and the lines disappeared, replaced by a solid blue screen. These events signaled that the subject could begin the movement. Once the subject started moving the lever, the movement was recorded; during this time the screen remained blue.

About 2 s after the movement, feedback was presented. This information consisted of the template for the task just attempted in white, superimposed upon the subject's movement trace in yellow. Both traces were again presented suddenly and simultaneously. The subject's trace extended beyond the end of the template to the right (it was sampled for 2 s), making it easily distinguishable from the template both in color and length. In addition, the root mean squared (RMS) error for that trial, calculated over the first 937 ms of the movement (the duration of the template), was presented together with the trial number above the traces. This feedback display remained on for 5 s, the screen was blanked for 2 s, and then the template for the next task to be attempted was displayed.

Reminder and no-reminder feedback conditions differed only in terms of the presentation just prior to each trial. In the no-reminder conditions, only the template for the task to be performed was presented. In the reminder conditions, the template had the subject's trace superimposed over it, exactly as it had been in the earlier feedback presentation. That is, the only difference between reminder and no-reminder conditions was that the reminder condition had the subject's trace superimposed over the template, with the RMS error above it, prior to each action, whereas the no-reminder condition did not.

Thus, the feedback originally received for the previous attempt at that task was repeated just before the subject's next attempt at the same task. Note that, for the blocked conditions, two identical feedback presentations were presented sequentially, with only 2 s separating them. But for the random conditions, the feedback report for the previous trial (of, say, Task A) would be followed by the template and reminder feedback for the next task (say, Task B). Note also that, for the random conditions, the reminder feedback and the performance to which it referred might be separated by as many as four trials of other tasks (and their feedback and reminder feedback reports); this was not the case for the blocked conditions.

After the tasks and procedures had been explained, subjects used Task C to perform three practice trials (see Figure 1), with feedback given as described earlier. Subjects' errors were discussed after each trial to be sure that they knew how to interpret the feedback.

**Design**

The experiment proper consisted of three phases: practice, immediate retention (given after a 5-min retention interval), and delayed retention (given on the next day). During practice, subjects performed 90 trials under one of four combinations of reminder feedback (present or absent) and spacing conditions (blocked or random), as described below. During the retention tests, subjects performed 12 trials (4 trials of each task in a random order), without feedback or reminder feedback.

Subjects were assigned randomly to one of four groups of 14 subjects each, organized in a $2 \times 2$ (Spacing $\times$ Reminder) factorial design. The random conditions had the three tasks in a quasi-randomized order, with the restriction no one task could appear on two consecutive trials. The blocked conditions had the tasks in blocks of 30 trials; all 30 trials of a given task were completed be-
before moving to the next task. For the blocked conditions, there were actually subgroups that differed with respect to the order of the three tasks: A, B, C; B, A, C; and C, A, B. The reminder and no-reminder conditions differed only with respect to whether or not the reminder feedback was presented prior to the next attempt at that task, as described earlier. Thus, the four groups formed were the blocked (B), random (R), blocked-reminder (B-rem), and random-reminder (R-rem) groups.

Performance Measures and Interpretations

To evaluate the role of spacing and reminder variations on learning, we focused mainly on performances in the retention tests. It is well known that learning effects can be obscured during the acquisition session because of the confounding of relatively permanent and temporary factors associated with the independent variables (e.g., Salmoni et al., 1984; Schmidt, 1988). Therefore, learning was measured on immediate (5-min) and delayed (1-day) no-KR retention tests, where conditions were equated for all groups (no feedback or reminders, with patterns given in a randomized order). The dependent variables of main interest are presented below.

RMS error. The main dependent measure was RMS error, a measure of overall proficiency that describes the extent to which the subject approximated the goal pattern on each trial (Schmidt, 1988). It was calculated from the start of the movement until the end of the template, or for 937 ms. Because the same RMS error can be produced by countless different movement trajectories, this measure is not particularly informative about the nature of any differences in overall performance.

Temporal and amplitude scaling of subject traces. To separate measures of how well the GMP and the time and amplitude parameters were produced, we developed a computer program that scaled (stretched or compressed) the subject-produced trajectory in time and amplitude so that the agreement with the goal pattern was maximized (see Wulf, Schmidt, & Deubel, 1993). The actual procedures for scaling the curves were as follows:

First, each subject-produced trajectory and the goal pattern were synchronized at the moment that a displacement of 1.5° from the baseline was achieved. In a few movements, however, small premovement instabilities (clearly not a part of the action) were taken by the computer as synchronization points; these cases could be easily identified, and the trial was resynchronized with our 1.5° criterion before further analysis. Synchronization at the start of the action ensured that the subject’s trajectory and the goal pattern were always aligned initially, which minimized errors caused by phase shifts from initial misalignment.

The experimenter-produced goal patterns were represented in the computer as position–time arrays with 187 individual values each. During data collection, each subject-produced trajectory was sampled at 200 Hz. Our procedure rescaled the subject traces in time, where the degree of rescaling could be varied with scaling factors ranging from .2 to 2.0. The value of the scaling factor indicated the proportion that the overall movement was stretched or compressed; a value of 2.0 implied a doubling of the movement time, for example.

Logically, the scaling process of the subject traces occurred as follows: For a scaling factor of 2.0 (for example), each sampled point (say, one occurring at 100 ms after the start of the action) would be mapped into another point delayed from the start of the movement by exactly twice as much (200 ms, for example), in effect doubling the duration of the rescaled movement and of all its parts. Next, the rescaled subject trace was smoothed and estimates were interpolated between the actual samples, such that the number of values during the goal movement time (937 ms) would be 187, the number of samples in the goal pattern over its movement time. The result was a rescaled subject trace with 187 samples and a duration of 937 ms. Next, the correlation between the goal trace and the rescaled subject trace was computed, using the 187 samples in each trace, with the correlation indicating the extent of agreement between the goal trace and the rescaled subject trace.

This procedure was repeated for each of the scaling values for scaling factors ranging from .2 to 2.0 (in steps of .1 units). The scaling factor that yielded the maximum correlation between the template and the rescaled subject trace was determined. To increase the precision of the estimate of the optimal scaling factor, we then repeated this procedure, using twenty .05-unit steps surrounding our initial estimate of the optimal scaling factor. We then selected the single scaling factor that produced the highest correlation between the goal trace and the subject trace. This scaling factor was taken as a measure of the error in temporal parameterization. In other words, this time-factor indicated the amount by which the subject’s trajectory had to be expanded or compressed in time so that it correlated maximally with the goal pattern. Values larger/smaller than 1.0 indicated that the movement was too slow/fast. Note that this process did not introduce any phase shifts between the two functions—only proportional (linear) scaling in duration, with the start of the movement fixed.

Next, an amplitude factor was determined as the variance in amplitude (computed over the 187 samples in each trial) of the optimally time-scaled movement trajectory, divided by the variance in amplitude of the goal pattern. This ratio indicated the proportion by which the subject trace had to be expanded (values less than 1.0) or contracted (values greater than 1.0) in amplitude so that it maximally approximated the goal trace. Thus, the amplitude factor was taken as a measure of error in amplitude parameterization.

Parameterization errors. As just described, the process of optimal rescaling of subject traces yielded what we have called a time factor and an amplitude factor. For each trial, these values are the scaling (or gain) factors
that describe how much the overall time or amplitude, respectively, had to be adjusted to generate the best fit with the goal pattern. The average (across trials, within subject) of these single-trial estimates of time and amplitude factors represents a constant error (CE) in time and amplitude parameterization (see Schmidt, 1988). The standard deviation (SD) of the time and amplitude factors across trials indicates the instability of the amplitude and time parameter assignment, a variable error (VE) measure. We interpret the amplitude and time factors and their VEs as measures of amplitude- and time-parameterization effectiveness, respectively.

Residual RMS error. After rescaling each subject trace in both time and amplitude, the remaining RMS error between the rescaled subject trace and the goal pattern was determined. We interpreted this residual RMS error to be a measure of the inaccuracy of the GMP, because it is sensitive to the disagreement between the subject trace and the goal pattern when the errors in amplitude and temporal parameterization have been removed.

Actually, two measures were derived from the residual RMS error calculations. First, the mean of the residual RMS errors over a series of trials gives an estimate of the deviation in the subject's average trajectory from the goal pattern. It is mathematically equivalent to measuring the average movement trajectory for a series of trials, and then computing the RMS deviation between this average trajectory and the goal pattern. In other words, the mean residual RMS error indicates the deviation (in RMS units) of the subject's average, or "typical," trajectory from the goal pattern. We take this measure to indicate the accuracy of the subject's GMP.

Second, we also wanted a measure sensitive to the stability of the GMP representation over trials, or the extent to which this representation was variable from attempt to attempt. Thus, what we term residual-RMS-error VE (a variation of variable error related to trial-to-trial variability) was calculated as follows, using again the example with 187 samples for the goal pattern: Within a block of several trials (e.g., 12 trials), and using the rescaled subject traces (with interpolated values), the within-subject (over trials) SDs of the amplitudes of the 1st sample, the 2nd sample, . . . , the 187th sample were computed. This procedure provided 187 separate estimates of the variability of the subject-produced templates. The mean of these 187 estimates provided a kind of "average" variability of the GMP and was the residual-RMS-error VE reported here. Of course, such a measure is insensitive to the fact that some portions of the movement are inherently more, or less, variable than others, and it tends to confound amplitude and temporal sources of variability. Also, this residual VE contains several additional sources of variability (recording errors, neuromuscular variability), so that this measure will surely not be a very pure estimate of the extent to which the GMP is unstable. But, within these limitations, we interpreted this as a measure that was at least sensitive to the instability of the amplitude characteristics of the GMP.

Results

Before turning to the main results, one fundamental assumption of the method for separating the parameterization and GMP errors should be examined. Before such an analysis could be interpreted appropriately, we had to verify that these movements were controlled by a single unit (or GMP). The theoretical basis of the rescaling procedures depended on the assumption that the entire movement was governed by a single, rescalable movement structure. If the action had been controlled by two or more programmed units—as were some of those in Young and Schmidt (1990) and Schneider and Schmidt (1993), who used a considerably different task—then scaling the movement proportionately in time and amplitude would not have been justified. At least a portion of the action (e.g., the interval between units) would not be expected to show proportionate scaling even though the movement time changed. We examined this aspect of the structure of the actions, using the methods of Young and Schmidt (1990).

Examining the Unit-Structure of the Patterns

The process involved computing the acceleration–time functions of the subject-produced trajectories and determining seven temporal landmarks defined by peaks, zero crossings, and valleys. These seven landmarks were labeled a, b, . . . , g from the earliest to the latest in the movement, and every landmark was present in every movement for every subject. The time of occurrence of each of these landmarks was measured by interactive graphics, and the within-subject (across-trial) correlations among these measures were computed. The analysis was based on the patterns of correlations between a–b, a–c, . . . , a–g, where the correlations were computed between the first landmark and successively later landmarks. We also examined the correlations a–g, b–g, . . . , f–g, that is, correlations of the various landmarks with the last landmark.

If these movements were represented by more than a single unit, or GMP, then there should be relatively high correlations (near 1.0, depending on noise, variability, and other factors) for all of the units for which the program is operating, but correlations should drop abruptly toward zero if the two landmarks should span the border of a unit. In one type of movement (termed Type II) studied by Young and Schmidt (1990; see also Schneider & Schmidt, in press), these correlations changed abruptly from about .80 (for a–d) to about .20 (for a–e), suggesting that the movement had more than a single unit whose boundary was between Landmark d and e. In some of their other actions (termed Type I), correlations changed gradually across the movement trajectories, with single steps yielding correlation changes that were always less than .15.
The patterns of correlations for Pattern A (using the last 10 trials of the practice phase) are shown in Figure 2, where the values are averaged across subjects (following $Z'$ transformation). The pattern of correlations for Patterns B and C did not differ from that for Pattern A in any important way, and they are not shown here. For all three movement templates, the patterns of correlation were consistent with the one-unit structure. As the separation between the landmarks increased from the start of the action (Figure 2, top, moving from a-b, a-c, . . . , a-g), the correlations always decreased gradually (from .85 to .56). With respect to the end of the action, as the separation between landmarks decreased (a-g, b-g, . . . , f-g), the correlations gradually increased (from .45 to .96; Figure 2, bottom). These separate interval-to-interval changes were always small; the largest for any interval in any task was .30, and most were smaller than .10. Further, this general pattern was present for each subject. Thus, we felt justified in using our scaling procedures here, as each of these actions appeared to be governed by a single GMP. See Young and Schmidt (1990; Schneider & Schmidt, in press) for more detail and rationale about this analysis.

Next, we present the results organized according to the three phases of the experiment: practice, immediate retention, and delayed retention. For each phase, the various measures of performance described in the previous section are discussed.

**Practice Phase**

**RMS Errors**

RMS errors during practice for the four groups are shown at the left of Figure 3. All groups reduced their deviations from the goal movement patterns across blocks. Performances of the B and B-rem groups were less consistent than those of the two R groups because the B and B-rem groups switched to different tasks before Blocks 3 and 5, with an associated elevation in error. The B and B-rem groups performed similarly across practice, whereas the R-rem group demonstrated larger error than the R group. Also, note that the magnitude of the decrease in RMS error across blocks is underestimated in Figure 3 because of the averaging of the relatively poor performance on Trial 1 (mean across groups = 717 units) with the subsequent trials in Block 1. There were no reliable differences among groups for these Trial-1 values; all $p$s > .05.
The analyses of variance (ANOVA) revealed significant main effects of reminders, \(F(1, 51) = 7.1, MSe = 388,011\), spacing, \(F(1, 51) = 24.3, MSe = 1,328,087\), and blocks, \(F(5, 255) = 40.0, MSe = 277,235\), all ps < .05. Also, the interactions of reminders and spacing, \(F(1, 51) = 5.8, MSe = 315,243\), blocks and spacing, \(F(5, 255) = 3.2, MSe = 22,303\), and blocks and reminders, \(F(5, 255) = 12.5, MSe = 86,660\), were significant, ps < .05. Overall, the two groups without reminders (R and B) performed with less error than the two groups with reminders (R-rem and B-rem), but the main effect of reminders occurred primarily because of the extremely poor performance of the R-rem group. Also, random practice produced generally larger errors than blocked practice. The interaction effects were mainly caused by the different performance profiles of the blocked and random conditions across blocks. In addition, the R-rem group had significantly larger RMS errors than the R group, \(F(1, 25) = 10.5, MSe = 65,846, p < .05\).

**Residual RMS Errors**

Residual RMS errors are our measure of the accuracy of the fundamental movement patterning, and are shown in Figure 4. The residual RMS error decreased across practice, \(F(5, 240) = 2.7, MSe = 0.04604, p < .05\), but there were no reliable effects of spacing, \(F(1, 48) = 1.6\), of reminders, \(F(1, 48) = 1.6\), or their interaction, \(F(1, 48) < 1\), all ps > .05. The residual RMS errors for the four groups tended to converge by the end of practice, but there were no reliable interactions between blocks and the other variables, all ps > .05. As with the RMS errors in Figure 3, the magnitude of the decrease in residual RMS error across blocks is underestimated in Figure 4 because of the averaging of the relatively poor performance on Trial 1 (mean across groups = 468 units) with the subsequent trials in Block 1. There were no reliable differences among groups for these Trial-1 values; all ps > .05.

**Variability of Residual RMS Error**

Figure 5 shows the variability of the residual RMS error, expressed as the within-subject mean of the one hundred and eighty seven 15-trial SDs, and interpreted as the instability of the GMP. This variability decreased significantly across practice, \(F(5, 250) = 5.0, MSe = 0.06098, p < .05\). In addition, there was a clear tendency for the reminder feedback to increase variability for both the random and blocked conditions. There was a significant effect of reminders, \(F(1, 50) = 10.6, MSe = 0.35207, p < .05\). Even during the practice phase, there was a tendency for the R group to have more stable movement patterns than the B group, but the effects of spacing, \(F(1, 50) = 2.4\), and the interaction of reminders and spacing, \(F(1, 50) < 1\), were not significant, ps > .05.
suggest that reminder feedback tended to increase the instability of the fundamental movement patterning throughout practice.

Amplitude and Time Factors

For the average values, all groups generally showed enhanced conformity to the overall temporal and spatial characteristics of the goal patterns as a result of practice (see Table 1). There was a significant block effect for both the time and amplitude factors, $F$s $(5, 240) = 11.6$ and $4.8$, $MSe$s $= 0.03026$ and $0.06708$, respectively, $p$s $<.05$. Generally, the R and R-rem groups demonstrated greater deviations in amplitude than the B and B-rem groups, $F(1, 48) = 5.8$, $MSe$ $= 0.35843$, $p < .05$, but the analogous effect for the temporal factor was not reliable. There was, however, a tendency for the R-rem group to have larger errors in time than the R group, whereas there was no effect of reminders in blocked practice; this was supported by the Spacing $\times$ Reminder interaction, $F(1, 48) = 4.4$, $MSe$ $= 0.02834$, $p < .05$. Thus, the presentation of reminder feedback appeared to decrease accuracy in the temporal factor (our measure of time-parameterization accuracy) in random practice but not in blocked practice.

The variability of the amplitude and time factors (see Table 1) were also reduced across practice, $F$s $(5, 240) = 13.4$ and $13.2$, $MSe$s $= 0.05386$ and $0.03301$, respectively, $p$s $< .05$. There was a significant Block $\times$ Spacing interaction for $VE$ of both amplitude and time factors, $F$s $(5, 240) = 3.9$ and $3.4$, $MSe$s $= 0.01565$ and $0.00845$, respectively, $p$s $< .05$, caused by the B and B-rem groups' having to switch tasks on Blocks 1, 3, and 5. There was a large effect of reminders, where $VE$s in both temporal and amplitude factors were larger (roughly 60% and 50% respectively) for the B-rem and R-rem groups as compared with the B and R groups, $F$s $(1, 48) = 12.3$ and $9.0$, respectively, $p$s $< .05$. There was no effect of spacing on either of these factors, nor were there Spacing $\times$ Reminder interactions; all $p$s $>.05$. Therefore, providing feedback reminders elevated the variability of parameter selection in practice. Interestingly, reminders even produced this effect in the blocked conditions, where the feedback and the reminder were presented sequentially between trials of the same task.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors in Parameterization for the Random, Blocked, Random-Reminder, and Blocked-Reminder Groups in Practice and Immediate and Delayed Retention</td>
</tr>
<tr>
<td>Practice blocks</td>
</tr>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Time factor CE (time parameter accuracy)</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>R-rem</td>
</tr>
<tr>
<td>B-rem</td>
</tr>
<tr>
<td>Time factor VE (time parameter stability)</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>R-rem</td>
</tr>
<tr>
<td>B-rem</td>
</tr>
<tr>
<td>Amplitude factor CE (amplitude parameter accuracy)</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>R-rem</td>
</tr>
<tr>
<td>B-rem</td>
</tr>
<tr>
<td>Amplitude factor VE (amplitude parameter stability)</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>R-rem</td>
</tr>
<tr>
<td>B-rem</td>
</tr>
</tbody>
</table>

*Note. R = random, B = blocked, R-rem = random reminder, and B-rem = blocked reminder group. Imm = immediate, and Del = delayed retention.*
Immediate Retention

RMS Error

The RMS errors for the immediate retention test (four trials of each task in a randomized order, without feedback or reminders) are shown in the right panel of Figure 3 (Imm). Group R was clearly more accurate than the other three groups, which appeared to perform similarly to each other. There was a significant effect of spacing, \( F(1, 51) = 9.0, MSe = 216,643, p < .05 \), whereas the effect of reminders, \( F(1, 51) = 3.3, MSe = 79,335.8, p = .08 \), reached only borderline significance. The interaction between reminder and spacing was significant, \( F(1, 51) = 5.4, MSe = 129,577, p < .05 \). Post hoc tests revealed that Group R was more accurate than Groups R-rem, B-rem, and B, which were not significantly different from each other. Importantly, the reminders degraded learning similarly to each other. There was a significant effect of spacing, reminders, \( F(1, 49) = 4.6, MSe = 0.00885, p < .05 \), as well as a significant Spacing \( \times \) Reminder interaction, \( F(1, 49) = 6.3, MSe = 0.01200, p < .05 \), but the main effect of spacing was not significant, \( p > .05 \). Post hoc tests indicated that the R group was significantly more accurate than all other groups, which did not differ significantly from each other. Thus, we interpret one effect of the reminder feedback as the degradation of motor-program learning when practice was under random conditions, but not under blocked conditions.

Residual RMS Error

The residual RMS errors showed effects similar to RMS errors (Figure 4). The residual RMS error was smaller for the R group (159 units) than for the B (202), R-rem (216), and B-rem (197) groups, which performed similarly to each other. There was a significant effect of reminders, \( F(1, 49) = 4.6, MSe = 0.00885, p < .05 \), as well as a significant Spacing \( \times \) Reminder interaction, \( F(1, 49) = 6.3, MSe = 0.01200, p < .05 \), but the main effect of spacing was not significant, \( p > .05 \). Post hoc tests indicated that the R group was significantly more accurate than all other groups, which did not differ significantly from each other. Thus, we interpret one effect of the reminder feedback as the degradation of motor-program learning when practice was under random conditions, but not under blocked conditions.

Residual-RMS-Error Variability

For the variability of the residual RMS errors—measures of the stability of the fundamental patterns (Figure 5)—we found a strong tendency for the groups with reminders (103 and 105 units for the R-rem and B-rem groups, respectively) to have larger variabilities than the groups without reminders (81 and 87 units for the R and B groups, respectively); this reminder effect reached borderline significance, \( F(1, 49) = 3.6, MSe = 0.05287, p = .06 \). There was no significant effect of spacing or of the Spacing \( \times \) Reminder interaction, \( ps > .05 \). Under both random and blocked practice, reminder feedback appeared to degrade the stability of the fundamental program for action.

Time and Amplitude Factors

For the CE\(s\) of the time factors, all groups performed essentially similarly (see Table 1), and there were no reliable effects of spacing, reminders, or their interaction, \( ps > .05 \). For the VE\(s\) of the time factors, the two reminder conditions (R-rem and B-rem) were 18% more variable than the two no-reminder conditions (R and B), but none of the main effects or interactions was reliable, all \( ps > .05 \).

For the CE\(s\) of the amplitude factors (see Table 1)—the measure of amplitude parameter accuracy—there was no reliable effect of either spacing or reminders. There was, however, a significant Spacing \( \times \) Reminder interaction, \( F(1, 49) = 5.2, MSe = 0.17562, p < .05 \); the R group (1.03) had a smaller amplitude factor than the B group (1.20), but the R-rem (1.16) and B-rem (1.09) groups did not differ reliably. All groups tended to make the movement too large, but this tendency was smaller for the R group than for the others. The VE of the amplitude factor was about 37% larger for the R-rem and B-rem than for the R and B conditions. This effect of reminders was significant, with \( F(1, 49) = 8.9, MSe = 0.02239, p < .05 \). The effects of spacing and the Reminder \( \times \) Spacing interaction were not reliable, however; \( ps > .05 \).

Delayed Retention

RMS Errors

The global RMS errors for the final retention test, seen to the right in Figure 3 (Del), showed a pattern very similar to that in the immediate retention test, except that all RMS errors were somewhat larger, possibly as a result of forgetting over the 1-day retention interval. The R and R-rem groups (557 and 645 units, respectively) tended to outperform the B and B-rem groups (692 and 671 units, respectively). This effect of spacing was reliable, \( F(1, 52) = 4.3, MSe = 180,996, p < .05 \). There was no effect of reminders, however, \( F(1, 52) < 1 \). The R group tended to outperform the R-rem group, but the B and B-rem groups performed similarly, suggesting an interaction between spacing and reminders (as found in the immediate test), but this interaction effect was not significant, \( F(1, 52) = 1.9, p > .05 \).

Residual RMS Errors

The residual RMS errors (Figure 4, right) were larger for the B and B-rem groups (309 and 247 units, respectively) than the R and R-rem groups (183 and 233 units, respectively). This spacing effect was significant, \( F(1, 51) = 4.4, MSe = 0.06760, p < .05 \), suggesting that random practice produced more accurate motor-program representations than blocked practice. There was a tendency for the reminder feedback to degrade learning in the random conditions, but not in the blocked conditions; however, there were no reliable effects of reminders, \( F(1, 51) < 1 \), or of the Reminder \( \times \) Spacing interaction, \( F(1, 51) = 2.9, MSe = 0.04412, p = .09 \).

We also ran pairwise comparisons between the R group and the B group, as well as between the R group
and the R-rem group, even though the Reminder × Spacing interaction was not quite significant, so these results should be viewed with some caution. However, because we are interested in (a) the effect of random versus blocked practice and (b) the effect of reminders versus no reminders on the learning of GMPs, these pairwise comparisons seemed interesting. This is especially so given the fact that the reminder feedback was simply repeated feedback (delayed only by a few seconds) in the blocked condition. In any event, the analyses indicated that the R group was significantly more accurate in residual RMS error than both the B group, \( F(1, 26) = 4.9, \text{MSE} = 0.11259, \) and the R-rem group, \( F(1, 25) = 4.8, \text{MSE} = 0.01725, p < .05. \) Thus, GMP learning was enhanced more by random than blocked practice, and reminders degraded program learning during random practice.

**Residual-RMS-Error Variability**

For this measure, the reminder feedback had particularly degrading effects in the random condition, with variability being 111 and 177 units for the R and R-rem groups, respectively. However, reminder feedback tended to decrease variability under blocked conditions (182 and 145 units for the B and B-rem groups, respectively). Despite these large differences, the Spacing × Remind interaction was not quite significant, \( F(1, 51) = 3.5, \text{MSE} = 0.03706, p = .07. \) The main effects of spacing and of reminders were not significant either, \( F(1, 51) < 1. \)

**Time and Amplitude Factors**

For the CEs of the time factors, the B group was surprisingly more accurate than the R, R-rem, and B-rem groups, which did not differ (see Table 1). The main effect of reminders was significant, \( F(1, 51) = 5.0, \text{MSE} = 0.05914, p < .05. \) The effect of spacing, \( F(1, 51) = 4.0, \text{MSE} = 0.04649, p = .05, \) and the interaction of spacing and reminders, \( F(1, 51) = 3.8, \text{MSE} = 0.04491, p = .06, \) were marginally significant. For the VEIs of the time factors, none of the main or interaction effects was reliable, \( p > .05. \)

For the CEs of the amplitude factors (see Table 1), none of the effects of spacing, reminders, or their interaction was significant, \( p > .05. \) For the VEIs of the amplitude factors, all groups performed at about the same level, and there were no reliable effects of either reminders or spacing, \( p > .05. \)

**Discussion**

Several sets of findings concern the role of reminder feedback and its differential effects as a function of the practice schedule. But before examining this major focus of the experiment, we turn briefly to a discussion of the nature of the performance differences found as a function of the type of practice.

**Effects of Random and Blocked Practice (Groups R and B)**

Considering only the groups without reminder feedback (Groups R and B), the RMS error was 25% and 17% smaller in the R group than in the B group in the immediate and delayed tests, respectively. Advantages of random over blocked practice have been shown many times previously, but the present results generalize the principles to somewhat more complicated tasks. More importantly, though, our analyses of the movement patterns provide evidence about the specific nature of the changes in movement control that occurred as a result of these different practice schedules.

The principal effects of random versus blocked practice were that the R group had a smaller residual RMS error—our measure of the GMP's accuracy—than the B group. These residuals were 21% smaller in the immediate test, and 41% smaller in the delayed test. We also found that the variability of this measure—interpreted as the instability of the subject's GMP representation—tended to be reduced by random practice, but these differences were not reliable for either retention test. The advantage for the R group probably did not result simply because we conducted their retention test under the same (randomized) conditions as we had given them in practice although we required the B group to switch conditions for the retention test. For example, Gabriele, Hall, and Buckolz (1987) and Hall, Domingues, and Cavazos (1994) have shown learning advantages for random over blocked practice even when the retention test was done under blocked conditions. We argue, that, relative to blocked practice, random practice resulted in the learning of GMPs that were more accurate representations of the target templates.

This was not the only source of the difference between random and blocked practice effects, however. In the immediate test, both groups made the movement too large, but the blocked group made the movement much larger than the random group did. There was also less variability in the amplitude parameter for the random group. These parameterization effects did not persist into the delayed retention test, however, and they were not present with the measures of temporal parameterization, so it is unclear how important these influences might be. It seems reasonable to conclude that the major source of benefit for random practice was more accurate GMPs. This finding has important implications for practice scheduling in practical situations, as changes in GMPs are regarded as far more difficult to learn and more important for skilled proficiency than are changes in parameterization (Schmidt, 1991b).

But perhaps more importantly, the findings have strong implications for the notion of GMPs in general. Several lines of new evidence have provided support for the idea that GMPs and their parameters are empirically separable—a notion that has been under considerable
debate for the past decade or so. In particular, we showed here that the blocked–random practice manipulation influences the acquisition of GMP accuracy but does not affect the learning of parameterization. Also, Wulf et al. (1993) showed that reduced feedback frequency facilitated the learning of GMPs but degraded the learning of parameterization. These dissociations not only provide support for the theoretical separation of GMP and parameterization processes, but they suggest that different principles are used to learn these two kinds of processes. Again, such findings could have important implications for practical applications where feedback and practice scheduling are of concern.

The Role of Reminder Feedback in Learning

The major purpose of the present experiment was to examine the effects of reminder feedback to test a feedback-induced-variability hypothesis for frequent-feedback effects. We separate the discussion of these effects according to their roles during the practice and retention phases of the experiment.

Effects in Practice

During the practice phase, the effects of reminder feedback on overall performance (i.e., RMS error, Figure 3) depended strongly on the conditions of practice—random or blocked. For RMS error, there was almost no effect of reminders for the blocked condition, which was not surprising in view of the fact that the feedback and reminder contained exactly the same information separated by only 2 s. But for the random condition, where reminders repeated feedback information that occurred from one to four trials earlier, reminder feedback degraded performance markedly, increasing RMS error by approximately 33% as compared with the no-reminder condition.

The analysis of the components of RMS error revealed that the effect was caused by elevated variability in numerous aspects of performance. The stability of the GMP representation (the residual-RMS-error VE), and the stability of both time and amplitude parameterization (VEs of the time and amplitude factors) were all degraded by reminder feedback. However, the accuracy of the GMP, indexed here by the residual RMS error, did not change as a function of reminder feedback. It seemed clear that reminder feedback induced variability in responding during random practice.

Reminder feedback represents additional information that might be used in organizing the subsequent trial. It informs the learner of the pattern error(s) on the previous attempt at that task, tends to compensate for forgetting of the feedback from the previous attempt, and should allow a more effective plan for correcting the next attempt. It is clear that subjects used this information but that such feedback use was severely detrimental to performance, not beneficial. These outcomes were not surprising in light of Nicholson's (1992; Nicholson & Schmidt, 1991) findings that feedback induced variability in responding, even when subjects were instructed to ignore it. Perhaps the additional feedback encouraged overcompensations for earlier errors, leading to highly variable performance both in terms of the GMP representation and its parameterization. Whatever the actual mechanism leading to the added variability, these findings provided a basis for using reminder feedback as a way to test the feedback-induced-variability hypothesis for frequent feedback.

Effects on Learning

The learning effects of reminder feedback again depended strongly on the practice conditions. For blocked practice, reminder feedback had almost no effect on overall performance (RMS error) in either retention test. For random practice, however, reminder feedback produced severe decrements in learning. In fact, reminder feedback reduced the learning in the random condition almost to the level of the blocked conditions, essentially obliterating the usually beneficial effects of random over blocked practice. This effect was almost completely caused by an elevation of residual RMS error, which was 36% larger for the R-rem group than for the R group in the immediate test and 27% larger in the delayed test. As before, we interpret these elevations in residual RMS error in terms of a systematically less accurate GMP. Here, the average trajectory for the R-rem subjects consistently deviated from the goal pattern more than for the R subjects. Reminders also tended to increase the variability of the residual RMS error—our measure of GMP stability—but these effects were generally not statistically reliable.

Finally, it was surprising in a general way that variability induced from reminder feedback, which, on the surface, would seem to be so "helpful" for correcting errors, would be so detrimental for learning. But these findings raise an apparent paradox about the role of variability. On the one hand, in so-called variability-in-practice studies (see Schmidt, 1988, chap. 14), intentionally varying parameters of a given GMP (which increases performance variability in practice) usually facilitated learning, at least as measured on novel transfer tests. On the other hand, feedback-induced variability during acquisition degraded learning in the present study. Variability during practice appears to have different roles in these two paradigms. One speculation is that, in the variability-in-practice literature, the learner is asked to produce different outcomes on different trials of the same task (e.g., shorter or longer movement times with the same GMP), which presumably requires different parameter values across trials; the subject learns the relationship between parameters and outcomes they produce (Schmidt, 1975). In the present study, however, the added variability from reminder feedback (and, we argue, from frequent feedback generally) is the result of the learner's attempting to produce the same outcome every time that task is re-
quested, while trying to compensate for random processes in movement production; the subject fails to learn the nonexistent relationship between the random noise and the required compensation, and GMP learning suffers as a result. It seems likely that the effect of variability on learning will depend on the way that variability is produced.

**Proactive Effects of Feedback**

These data raise several possibilities about the mechanisms by which frequent feedback degrades learning (see Schmidt, 1991a). In previous work and theorizing, the major focus has been on what may be termed the retroactive effects of feedback, related to how feedback is linked or associated with the previous response to which it applies. In the present experiment, however, we have shown a very different example of frequent feedback that degrades learning, but here it was provided by feedback reminders. The most probable way that reminder feedback operated here was proactively—not retroactively—as its major effects would seem to have been on processes associated with the upcoming trial that followed in a few seconds without interruption. On the contrary, any proactive operations must have spanned from one to as many as four other earlier trials (and their feedback and reminder feedback presentations). Although such proactive processes were not ruled out here, it is difficult to imagine how reminder feedback could have operated to degrade learning when the two sources of information (the reminder and the trial to which it referred) were so severely separated.

These findings thus support the hypothesis that a source of the decrement in learning that is caused by frequent feedback in practice is related to its variability-producing effects, regarded as maladaptive short-term corrections, that degrade learning. Other versions of a general proactive view are not eliminated by these results, however, such as the notion that frequent feedback, in addition to generating excessive response variability during acquisition, also blocks subjective retrieval operations, as discussed next.

**Blocking Retrieval Operations**

A second proactive hypothesis holds that frequent feedback—and, in the present experiment, reminder feedback—provides excessive guidance for the production of the subsequent trial. Thus guidance could block subjective retrieval operations that have been found to be important for learning in other situations (Bjork, 1975; Magill & Hall, 1990; Schmidt & Bjork, 1992). Prior to action, the subject must presumably retrieve the GMP for the pattern to be produced on that trial and parameterize it, both of which might be effectively obviated by providing reminder feedback. Against this view is the finding that reminder feedback did not facilitate performance in acquisition as one would have expected if the reminders were serving a purely retrieval function. Rather, reminder feedback degraded performance in acquisition for random practice, mainly by producing excessive response variability.

Even so, the reminders might have two simultaneous roles. First, they could induce variability (as shown during acquisition here), worsen performance in acquisition, and degrade learning by disrupting response stability. But, second, they could also be providing cues that block the subject's need for retrieval operations, which degrades learning. That is, even though the performance is being disrupted by excessive variability, the learner could still be guided through the retrieval operations by the directive properties of reminder feedback. Both processes could be operating to degrade learning here.

The finding that reminder feedback degrades learning in random practice but not in blocked practice supports this retrieval hypothesis. In blocked practice, the tendency for subjective retrieval operations is already reduced by the fact that the previous trial involved the same task and parameters, allowing the learner to retrieve a new GMP and having to parameterize it before each trial. Therefore, any retrieval cues provided by reminder feedback in blocked practice are ineffective in providing further reductions in retrieval practice. However, in random practice without reminders, retrieval operations are typically required because the learner must clear working memory of the GMP from the previous trial and retrieve and parameterize another GMP for the next trial. After working memory has been cleared, presenting reminder feedback could substitute for the subjective retrieval operations required without reminder feedback. In this sense, giving reminder feedback would seem to weaken the beneficial effects of random practice. This is essentially what was found: Reminder feedback eliminated the learning benefits of random over blocked practice.

It is even possible that the learning processes underlying reminder feedback, on the one hand, and the processes underlying blocked practice, on the other hand, are fundamentally the same, with both of them blocking retrieval operations. One line of support for this contention is that reminder feedback had the same general effect on learning—reduction in the accuracy of the GMP—as blocked practice did, and neither affected our measures of parameterization learning. We may even view this hypothesis in another, more extreme way, wherein random practice could be more effective for learning than blocked because of the tendency for feedback to be difficult to use when tasks are randomized, rather than because of a benefit of randomized scheduling per se. Thus, random—blocked effects might really be feedback-use effects, in the final analysis.

Future work will be needed to decide whether reminder feedback degrades learning because of its variability-inducing effects, or because it operates like blocked practice to obviate retrieval practice. On balance, though, the present evidence showing that remind-
ers elevated variability in numerous aspects of performance during acquisition strongly favors the view that frequent feedback degrades learning at least in part because it produces excessive response variability in acquisition. These maladaptive short-term corrections apparently degrade the subject’s capabilities to acquire accurate GMPs for long-term use.

NOTES

2. The data from one subject in the random group on Day 1 were lost.
3. We do not report the absolute value of CE (i.e., |CE|) here because its effects tended to duplicate those of CE, and the CE measure is a better descriptor of the directional biases.

REFERENCES


Submitted January 7, 1993
Revised August 16, 1993